The Influence of Cast Shadows on Visual Search

Ronald A. Rensink
Departments of Psychology and Computer Science
University of British Columbia
Vancouver BC   Canada

Patrick Cavanagh
Department of Psychology
Harvard University
Cambridge MA   USA

In press, Perception. Do not quote or cite without permission.

Word count (abstract): 139

Word count (body): 6240

Author Notes

Much of this work was done while the first author was at Cambridge Basic Research, Nissan Research & Development, Inc., Cambridge, Massachusetts, USA. The authors would like to thank Jennifer Corbett, Carolyn Harp, Kristen Macuga, and Andrew Synyshyn for their help in running the experiments. Correspondence may be addressed to R.A. Rensink, Department of Psychology, 2136 West Mall, University of British Columbia, Vancouver BC V6T 1Z4, Canada. (Email: rensink@psych.ubc.ca).

Abstract

We show that cast shadows can have a significant influence on the speed of visual search. In particular, we find that search based on the shape of a region is relatively slow when the region is darker than the background and corresponds to a shadow formed via lighting from above. This pattern of results supports the proposal that an early-level system rapidly identifies regions as shadows and discounts them, making their borders more difficult to access. The constraints used by this rapid system are mapped out in regards to the luminance and texture of the shadow region, and the nature of the item casting the shadow. Among other things, this system is found to distinguish between line elements (items containing only lines) and surface elements (items containing visible surfaces), with only the latter deemed capable of casting a shadow.

The Influence of Cast Shadows on Visual Search

Introduction

One of the more fundamental dichotomies of visual processing distinguishes between operations carried out rapidly and in parallel across the visual field, and operations carried out sequentially on an item-by-item basis (Al Haytham, c.1000 / 1989; Helmholtz, 1867 / 1962; Neisser, 1967). In many current views of human perception, this division takes the form of an *early* stage that rapidly extracts *features* from the incoming image, and a subsequent *attentional* stage that assembles these features into more coherent descriptions of the external world (e.g. Beck, 1982; Julesz, 1984).

Although early-level features were originally believed to describe relatively simple aspects of the image (such as color and two-dimensional orientation in the plane), more recent work has shown that they can also describe more complex aspects of the three-dimensional (3D) scene, such as surface convexity (Ramachandran, 1988; Kleffner & Ramachandran, 1992), direction of lighting (Enns & Rensink, 1990), and occluded surface structure (Rensink & Enns, 1998). These results support the view that early vision attempts to extract as much scene structure as possible, and that even when a complete and accurate interpretation of the image is not possible, it is still worth forming a "quick and dirty" initial estimate (see Rensink & Enns, 1995, 1998).

The issue examined here is whether cast shadows are an aspect of the world that is also handled by such rapid processes at early levels. Cast shadows do not correspond to structure in the world itself, arising instead from interactions between the world and the sources that illuminate it (see, e.g. Baxandall, 1995). This distinction is an important one: If, for example, the lighting-dependent border of a shadow were mistaken for the lighting-independent border of an object, serious errors in scene interpretation could arise (Braje, Kersten, Tarr, & Troje, 1998; Cavanagh & Leclerc, 1989). It would therefore be useful for the visual system to identify regions as shadows as soon as possible, and then—after any relevant information has been extracted from them—discount these regions so that they do not interfere with subsequent processing (Figure 1). On the other hand, because a shadow always gives rise to a region in the image that can be readily interpreted as something else (e.g., a black patch on a surface), shadow interpretation is fraught with difficulties, especially if attempted in the absence of high-level knowledge of the scene (Cavanagh, 1991).

Insert Figure 1 about here

A few studies have begun to examine the issue of rapid shadow interpretation. Rensink & Cavanagh (1993) found that an “anomalous” (i.e., upside-down) shadow among a set of “normal” (i.e. upright) shadows was detected more quickly than a normal among anomalous shadows. They explained this in terms of a rapid system that discounted normal shadows, and treated the anomalous shadows as non-discounted regions. Cunningham, Beck, and Mingolla (1996) showed that shadows could rapidly provide an assignment of 3D depth to items with a cast shadow. Elder, Trithart, Pintilie, & MacLean (1998) found that cast shadows could be rapidly distinguished from attached ones, and that surface relief could be rapidly computed via shadow displacement.

In what follows we look at the issue of rapid shadow interpretation in greater detail, and develop a more robust methodology that can test for its existence. This methodology is based on visual search experiments involving simple figures that can correspond to items casting shadows. The approach was similar to that of Rensink & Cavanagh (1993), but more direct: Observers were asked to determine the presence or absence of a target item containing a region that differed in shape from the corresponding regions of the distractor items (Figure 2). Depending on the stimulus conditions, the region may or may not be interpreted as a shadow; search results can then be used to determine both the conditions that lead to shadow interpretation as well as the kinds of objects that can cast them.

Insert Figure 2 about here

Our results show that the shape of a region in the image is relatively difficult to access when it is identified as a shadow, and that this occurs exactly for those conditions compatible with the existence of shadows. This pattern of results supports the proposal that there exists a system in early visual that can rapidly identify shadows and then discount them in some way, even when such discounting interferes with search. This rapid shadow analysis does not seem to occur when the same displays are inverted, suggesting that the analysis relies on an assumption that lighting is from above. We also find that for these conditions the ratios of target-absent to target-present slopes are much higher than for most search tasks, suggesting that shadows may also affect the strategic control of the search process.

General Design

Methodology

The experiments described here were based on the same methodology as for most visual-search experiments, in which the observer is asked to search a display as rapidly as possible for an item containing a unique visual attribute (a *target*) among a set of other items (*distractors*). A target was present on half the trials (chosen randomly) and absent on the remainder. Observers were asked to determine the presence or absence of the target item as quickly as possible, while keeping errors below 10%. Reaction time (*RT*) was measured as a function of the number of items present in the display (*set size*). The primary dependent variable was *search rate*, defined as the slope of the correct RT over set size. (For more detailed discussions of visual search methodology, see e.g., Treisman & Gormican, 1988).

For all experiments here, displays consisted of 2, 6, or 10 items. For each display these were positioned randomly on an imaginary 5 x 4 grid of possible locations, with locations selected such that the density of items was similar for all set sizes. The display area subtended approximately 10 x 8° of visual angle, with each item less than 2° in size. The position of each item was jittered by ±0.5° to minimize any possible effects of item collinearity.

Observers

Each experimental condition tested 10 adult observers with normal or corrected-to-normal visual acuity. About half the observers were naive to visual search methodology, while the others had been tested extensively in other search (or related) tasks. All observers completed 4 sets of 60 trials in each condition of each experiment.

Stimuli

Stimuli were vertically-oriented rectangles, each with an attached quadrilateral that could—at least under some conditions—be interpreted as a shadow cast by a vertical post (i.e., the rectangle) onto a ground plane (Figure 2). Rectangles were usually of a uniform color (medium gray), and were usually outlined (in white) to separate them from the background (also medium gray). Targets were usually items with quadrilaterals extending at an angle of 30° from the horizontal (in the image plane); distractors had quadrilaterals at 60°. Areas of all quadrilaterals were adjusted to be approximately equal, so that observers could not base search on their area, but required access to their shape.

Each observer was tested on two counterbalanced conditions: *upright*, in which the quadrilaterals corresponded to shadows created via lighting from above (Figure 2a), and *inverted*, where items were rotated by 180° in the image plane so that quadrilaterals corresponded to shadows created via lighting from below (Figure 2b).[1] Since these differed by a 180° rotation, image properties were the same in both cases, while (inferred) direction of lighting differed.

Analysis

Data analysis was based on the *combined* rate for each condition, i.e., the average of the target-absent and target-present slope for each observer, weighted such that both components had the same average value in each condition. Detecting differences in speed between upright and inverted cases was based on paired t-tests (two-tailed) on the average target-present rates for each observer in a condition.[2]

Experiment 1: Basic Effect

For Experiment 1, items were based on the set of *standard stimuli*: targets and distractors both had gray quadrilaterals that were darker than the background and with sharp borders, corresponding to “hard” shadows; targets were distinguished by a unique two-dimensional orientation of the quadrilaterals (Figure 3). Search was then compared for upright and inverted conditions.

\-------------------------------------
Insert Figure 3 about here
\-------------------------------------

If shadow identification does not occur at early levels, targets and distractors would be distinguished only by the two-dimensional orientation of the quadrilaterals, which would be the same for upright and inverted. No differences in search speeds should therefore exist between the upright and inverted cases. But if rapid shadow interpretation exists, it may use some of the same assumptions as other rapid-interpretation processes; in particular, it may assume that lighting is from above (Enns & Rensink, 1990; Kleffner & Ramachandran, 1992). Shadow interpretation would then occur for the upright but not the inverted case, leading to a difference in search speeds.

For Condition 1A, stimuli were exactly the standard set, which could be interpreted as a set of posts casting shadows onto a ground plane (Figure 2a). Results are shown in Figure

3. Search for the upright case was 13.4 ms/item for target-present trials, and 40.3 ms/item for target absent trials. Meanwhile, search for the inverted case was reliably faster: speeds were 4.8 ms/item for target-present trials, and 22.4 ms/item for target-absent trials.

To show that this difference in speed was not somehow due to the shape of the items, Condition 1B reversed the intensities so that light regions were now dark, and dark regions now light (Figure 4). The difference in search rates now disappeared, with speed for upright items (5.1 and 12.1 ms/item) no different than for inverted (5.1 and 11.4 ms/item). Absolute speeds for both were similar to those of the inverted case of 1A, indicating a rapid search for distinctive image features in these three cases, with the difference in 1A due to a slowdown of search for upright. These results support the view that regions are not considered to be shadows if they are lighter than the surrounding background, and so are not discounted.

Insert Figure 4 about here

Although a cast shadow in the scene must correspond to a dark region in the image, the converse does not hold: A dark region can correspond a number of things in the scene, such as a surface of low reflectance, or a nonlambertian surface oriented obliquely to the direction of lighting (Baxandall, 1995). To test whether it was the dark quadrilaterals in Condition 1A that somehow caused slowdown, Condition 1C used the quadrilaterals alone, without accompanying rectangles. No slowdown was detected, with search for upright (1.9 and 7.7 ms/item) not differing from inverted (1.4 and 11.5 ms/item). This suggests that the relevant process may need a nearby item to correspond to the *shadowcaster*—the object in the world that casts the shadow[3].

To further examine whether early vision is sensitive to shadows or just to dark patches, Condition 1D restored the rectangles, but now surrounded the quadrilaterals of the standard stimuli with white borders. If slowdown is due to a system that handles dark patches, there should be no large effect. But if it is due to a system specialized for shadows, slowdown will likely be eliminated, since white borders do not generally surround shadows. Speeds were similar for upright (5.5 and 11.2 ms / item) and inverted (4.7 and 10.2 ms / item). Slowdown therefore appears only in conditions compatible with the identification of a region as a shadow.

Condition 1E examined whether a similar effect occurs if borders are black rather than white (Figure 4). Results show that this did occur, with similar speeds for upright (3.7 and 10.8 ms / item) and inverted (2.4 and 7.5 ms / item). Evidently, it is not sufficient that the region beside a shadowcaster be darker than the background—a step in the luminance profile at the border will prevent slowdown, even if the region is dark. This is consistent with reports based on subjective impression, where a dark surrounding line will prevent the description of a region as a shadow (Hering, 1874 / 1964).

To investigate whether this absence of slowdown stems from the sharpness of the internal luminance profile of the border, or from the fact that the interior of the region is lighter than the border, Condition 1F used stimuli which had a gradual luminance lightening inside the quadrilaterals (Figure 4). Slowdown here was again absent, with no difference in slopes for upright (2.4 and 12.6 ms / item) and inverted (2.1 and 10.0 ms / item). As such, slowdown appears to be absent whenever areas inside the region are lighter than the border. This is a natural consequence of a shadow interpretation system.[4]

Condition 1G tested whether the sharpness of the luminance profile along the border has an effect. Here, a narrow black region was inserted inside each quadrilateral, resulting in a monotonic decrease in darkness towards the edges of the region (Figure 4). No differences in speed between upright (3.2 and 15.4 ms / item) and inverted (3.0 and 17.4 ms / item) were found. Thus, a sudden change in luminance profile is also sufficient to prevent slowdown.

To determine if any nonuniformity of the luminance profile of the border is enough to prevent slowdown, Condition 1H blurred the edges of the standard quadrilaterals so that sharp edges were removed (Figure 4). Slowdown now returned, with search for upright (14.3 and 42.9 ms / item) reliably slower than for inverted (5.5 and 22.0 ms / item). Such a pattern would natural for a shadow-interpretation system: blurred edges are part of “soft” shadows, which occur when penumbras are formed via diffuse or extended light sources (e.g., Baxandall, 1995).

Taken together, the results of Experiment 1 provide strong evidence of a slowdown in search for dark regions that can be interpreted as cast shadows due to lighting from above. This slowdown was not found for conditions incompatible with shadows, such as regions lighter than the background, or with sharp luminance profiles along their borders. This pattern of results is difficult to explain based on image properties—why should a simple 180° rotation in the image plane create such effects? But it is compatible with the quadrilaterals being rapidly identified as shadows (based on an assumption of lighting from above) and then discounted in some way.

Experiment 2: Alternate Approaches

To further test whether the slowdown effect found in Experiment 1 is indeed due to a discounting of shadows, Experiment 2 used an alternate approach in which the target was defined by a combination (or *conjunction*) of features instead of a feature *difference*. Condition 2A attempted to exploit discounting to *speed* search, not slow it. The target here was that of Condition 1D (quadrilaterals with white borders), while distractors were of two kinds: (i) those of Condition 1D, and (ii) the targets with the white borders removed (see Figure 5). Targets are therefore characterized by a conjunction of dark region (non-discounted non-shadow) plus orientation of 30°.

Insert Figure 5 about here

If shadows are rapidly identified and discounted, search for upright should be faster than for inverted, since one of the relevant distractor properties (the 30° orientation) is now less accessible and so would interfere less. In this type of search, then, discounting would cause a speedup rather than a slowdown. Results show that such speedup does indeed occur: search was faster for upright (30.4 ms and 68.2 ms / item) than for inverted (41.4 ms and 96.0 ms / item).

Condition 2B examined a variant in which the quadrilateral borders were switched—i.e., quadrilaterals with borders in Condition 2A no longer had them, while quadrilaterals originally without borders now did (Figure 5). The target is therefore defined by a conjunction of dark patch (without borders) plus orientation of 30°. If shadow interpretation occurs, targets will be distinguished by a property of their shadows, and so all shadows must be checked. Since shadows are relatively difficult to access in the upright case, this will result in a slowdown for the upright condition. Search for upright (54.5 and 109.3 ms / item) tended to be slower than for inverted (47.3 and 87.1 ms / item), although this difference was not reliable.

To examine whether reliable slowdown could be achieved with a different set of stimuli, Condition 2C replaced the distractors containing the 30° quadrilaterals with rectangles at a 30° orientation in the image plane. As in the case of dark areas surrounded by white lines, these isolated dark areas are not interpreted as shadows (Condition 1C). Targets were therefore again defined by a conjunction of dark non-shadow plus an orientation of 30°. Search for upright (32.5 and 53.1 ms / item) was now reliably slower than for inverted (20.6 and 38.2 ms / item), providing further evidence that regions interpreted as cast shadows are discounted, with their properties being less accessible to higher-level processes.

Experiment 3: Textured regions

Given that the techniques developed in Experiments 1 and 2 can determine when rapid shadow interpretation occurs, it becomes possible to map out the sensitivity of this system to various aspects of the stimuli. As a test of this, Condition 3A examined the sensitivity of this system to the texture of a region. A set of thin white stripes was placed on the quadrilaterals of the standard stimuli, so that they now were textured (Figure 6). This eliminated slowdown: search for upright (4.9 and 18.6 ms/item) was no slower than for inverted (9.2 and 19.3 ms/item), suggesting that regions with stripes are not interpreted as shadows.

Insert Figure 6 about here

However, it is worth considering the possibility that this absence of slowdown was because observers could attend to parts of the items (by accentuating the quadrilaterals or inhibiting the rectangles), making the quadrilaterals the only relevant items in the display. Such guided search can be done for color (Treisman & Gelade, 1980), and it may be that this could also be achieved on the basis of texture. To test this possibility, Condition 3B used items with the same texture in both the quadrilaterals and in the rectangles (Figure 6), so that selection via texture could not be done. Search here did not differ reliably for upright (4.1 and 23.9 /item) and inverted (5.6 and 20.2 ms/item), supporting the proposal that the lack of slowdown in 3A was indeed due to a lack of shadow interpretation.

To investigate whether the presence of texture in the quadrilaterals is the key factor in the absence of slowdown, Condition 3C placed stripes on the background only, leaving the quadrilaterals as solid dark regions. Slowdown was again absent: slopes for upright (4.3 and 16.6 ms / item) did not reliably differ than for inverted (2.6 and 9.5 ms / item); for these stimuli, the dark patch may have been interpreted as a solid structure that obscured the texture of the background. In any event, this result indicates that the key factor preventing shadow interpretation is not the presence of a texture within the region—rather, it is the presence of a texture *boundary* aligned with its border.

To test this further, Condition 3D placed stripes in both the quadrilaterals *and* the background. Slowdown now returned, with search for upright (12.5 and 34.9 ms / item) reliably different than for inverted (4.4 and 18.1 ms / item). This again indicates that the factor preventing shadow interpretation is a texture boundary aligned with the border, and not the simple presence of texture *per se*. If such a boundary exists, the interior is treated as a solid structure; otherwise, the texture is seen as extending through the background, with the dark region interpreted as a shadow falling upon it.

To determine if these effects were due to virtual lines formed by the terminators of the stripes, or to contrast reversals along the texture boundary (Cavanagh & Leclerc, 1989), Conditions 3E-H used similar conditions with a texture formed by a set of white stipples which had minimal contact with the border of the region (Figure 6). Condition 3E corresponded to Condition 3A, where the texture was restricted to the quadrilaterals.

Consistent with Condition 3A, speed for upright (13.9 and 39.6ms / item) was no slower than for inverted (12.0 ms and 35.9 ms / item).

The hypothesis that search could be guided to the textured regions was again tested in Condition 3F, where both the rectangles and quadrilaterals were filled with the same texture. Again, no reliable differences were found between upright (7.4 and 19.9 ms / item) and inverted (7.2 and 14.5 ms / item), suggesting that the rapid search did not depend on texture-guided access that isolated the quadrilaterals.

Condition 3G paralleled Condition 3C in that the background alone was textured. As before, slowdown did not occur: speed for upright (3.5 and 19.6 ms / item) was not reliably different than for inverted (4.4 and 13.3 ms / item).

Finally, Condition 3H used a lined texture that ran through both the background and the quadrilaterals. As in Condition 3D, a slowdown was found: search for upright (12.0 and 34.3 ms / item) was reliably slower than for inverted (6.5 and 22.2 ms / item). Consequently, both kinds of textures yield the same pattern of results: a dark region will not be interpreted as a shadow if a texture boundary is aligned with its border.

Experiment 4: Textured borders

Experiment 4 investigated whether a texture boundary along the border of a dark region will prevent shadow interpretation even for the case of a textured (or patterned) line, where no textures exist in the region or its surround. Condition 4A placed dashed lines along the borders of the quadrilaterals (Figure 7). Slopes for upright (6.9 and 17.2 ms / item) did not reliably differ from those for inverted (7.7 and 18.8 ms/item). As such, it appears that the texture boundary need not be continuous to prevent shadow interpretation.

Insert Figure 7 about here

Condition 4B investigated this further by placing two white dots along the border at the far corners of the quadrilateral (Figure 7). Again, slopes for upright (3.2 and 17.2 ms/item) did not differ from those for inverted (3.0 and 17.1 ms/item). Even a pair of dots along the border appears to be enough to prevent shadow interpretation.

Condition 4C removed one of the dots, so that only one dot remained at the far corner. This again was enough to prevent slowdown: speed for the upright case (6.5 and 18.7 ms/item) was not reliably different from that of the inverted case (6.1 and 11.8 ms/item).

Condition 4D tested whether the precise position of the dot along the border was important, placing the dot midway along the far luminance boundary (Figure 7). Although a trace of a slowdown was found, it was still largely absent: slopes for upright (4.1 and 22.3 ms/item) did not reliably differ from those for inverted (1.5 and 17.1 ms/item). Evidently, even a single dot located anywhere along a luminance boundary can inhibit rapid shadow interpretation.

All results of Experiments 3 and 4 are therefore consistent with the proposal that the rapid interpretation of a region as a shadow requires it be darker than its surroundings, and that no texture boundaries or small elements be aligned with its (luminance) border.

Experiment 5: Shadowcaster Constraints

Another interesting issue that can be investigated using these techniques concerns the structure of the shadowcaster. All standard stimuli contain a rectangle formed of a medium gray region surrounded by a white line (Figure 2). Since the background is also medium gray, this can be interpreted as a *surface element* (an element with an opaque surface assigned to it) or as a *line element* (an element with no assigned surface structure); more concretely, the outline could correspond to a rectangular surface, or a wire loop. It is worth testing which of these interpretations is used: Given the need for a shadowcaster in rapid interpretation, it may be important to establish that it corresponds to a surface element: A line element would not be able to cast a noticeable shadow.[5]

In Condition 5A, a gap was placed in the line surrounding the rectangles (Figure 8), so that the outline had to be interpreted as a line element. Interestingly, no difference was now found between speeds for upright (6.0 and 18.4 ms/item) and inverted (5.1 and 16.1 ms/item). This indicates that a distinction between line elements and surface elements does exist, and has consequences for shadow interpretation: a region will be interpreted as a shadow only if the shadowcaster is a surface element.

Insert Figure 8 about here

Condition 5B examined this further by replacing the rectangular outline by a simple vertical line. Again, no difference in search speed was found between upright (1.9 and 14.7 ms/item) and inverted (3.7 and 9.6 ms/item).

Condition 5C investigated whether the shadowcaster could be considered a surface element if a texture were in its interior (Figure 8). Results show little evidence of slowdown, with search for upright (8.8 and 33.8 ms / item) being marginally slower than for inverted (4.0 and 17.0 ms / item).

Condition 5D investigated whether the effect of gaps could be “repaired” by grouping. Here, gaps were placed in the top and bottom of the outline, creating a pair of brackets that could be easily grouped (Figure 8). However, search for upright (3.3 and 22.7 ms / item) was no faster than for inverted (5.7 and 15.2 ms / item).

Condition 5E tested whether status as a surface element could be restored if the gap were part of a virtual line formed by a series of dots (Figure 8). Again, no reliable difference was found between the upright (3.8 and 20.7 ms / item) and inverted (3.1 and 12.4 ms / item).

Finally, Condition 5F examined whether virtual lines formed by texture boundaries were sufficient to establish the shadowcaster as a surface element. Slowdown was again absent, with no reliable difference in speed between the upright (6.1 and 47.2 ms / item) and inverted (4.5 and 22.5 ms / item) cases.

The results of this experiment therefore indicate that at least two distinct kinds of structure exist at early levels: line elements (with no surface structure) and surface elements (with an opaque surface structure). Rapid shadow interpretation is possible only when the shadowcaster is a surface element. Such a surface element appears to require the presence of a continuous luminance border around it, or possibly texture in its interior. Virtual contours and grouping appear to be insufficient to establish status as a surface element.

General Discussion

The experiments presented here show that visual search based on the shape of a region is slowed down whenever that region corresponds to a cast shadow falling on a ground plane. Regions not interpreted as shadows give rise to search that is extremely fast, typically just a few milliseconds per item. As such, the process responsible for the slowdown must act preemptively, prior to this rapid search, acting very rapidly and in parallel—characteristics generally taken to indicate processing at early levels of vision.

This pattern of results supports the existence of a process at early levels that can rapidly identify regions as shadows and then discount them to some extent, making the shapes of the discounted regions relatively difficult to access, even when this discounting interferes with the task. If not interpreted as a shadow, the region is not discounted and so remains available for rapid search. In most of the conditions tested here, this difficulty of access caused search to slow down; however, it is also possible to set up situations where it causes search to speed up (Condition 2A). It is worth mentioning that the discounting of shadows may not be complete, in that the magnitude of the slowdowns are not what might be expected were the distinctive areas completely inaccessible to search. Such "partial" discounting may explain why cast shadows can sometimes interfere with recognition (Braje et al., 1998).

In any event, these effects are found for dark regions but not light ones, and are destroyed if any texture (or pattern) boundary is aligned with the border of the region, or if the region is not near an item with a visible area. They therefore support a clear distinction between the rapid handling of shadows and two related processes: the rapid handling of reflectance (e.g Mitsudo, 2002), and the rapid handling of shading (e.g., Kleffner & Ramachandran, 1992), neither of which have (nor should have) this particular set of constraints.

Rapid Interpretation Process

The pattern of slowdown found in these experiments has a consistency that points to a small number of constraints that must apply to the region under consideration if it is to be considered a shadow. As expected of a rapid interpretation process (see Rensink & Enns, 1995, 1998), these constraints take the form of rules that are simple and can be quickly applied in parallel at each location in the image:

a. the border of the region must be darker than its surroundings (Condition 1B)
b. it cannot completely contain a lighter area (Condition 1E)
c. the border of the region cannot be aligned with a luminance or pattern boundary (Conditions 1D-E, Experiment 3)
d. the border cannot fall exactly on a small structure in the image (Conditions 4B-D)
e. the border cannot have a luminance profile with a sharp inner step (Condition 1G)
f. there must be a nearby element deemed to be the shadowcaster (Condition 1C)
g. the item corresponding to the shadowcaster must be a surface element, not a line element (Experiment 5)
h. lighting must be from above the observer, either left or right (Conditions 1A, 8C, inverted cases in all conditions)

These rules are based on constraints that clearly pertain to shadows. Because early-level processes do not operate using knowledge of particular objects, it is often impossible for them to determine whether a dark region in the image corresponds to a shadow, or to a dark patch or oblique surface in the scene (Cavanagh, 1991). Thus, early interpretation of a region as a shadow can never be guaranteed. But the set of constraints used by the rapid system appear to minimize the probability of the two types of error (viz., failing to identify a shadow as a shadow, and incorrectly identifying a nonshadow as a shadow), while still being able to identify cast shadows under a large variety of conditions.

The constraints used by the rapid system are far from arbitrary. To begin with, the requirements that the border and interior of the region be darker than its surroundings (a and b) are constraints well-grounded in the physics of shadow formation (Baxandall, 1995).

The prohibitions against alignment (c and d) result from a different set of considerations, viz., the low likelihood that a shadow would fall exactly on a texture boundary, or on a small element. In the case of (c), it is much more likely that a luminance border aligned with a texture boundary corresponds to a border between two different surface structures than to a shadow border that just happens to fall on a texture boundary. Similar considerations of likelihood also apply to (d), where it is unlikely that a shadow border just happens to fall on a small element on the ground. In both cases, the rapid system is somewhat conservative; for example, the constraint against texture boundaries does not apply to the use of texture for 3D shape, being replaced by a more liberal constraint of no contrast reversals along the border (Cavanagh & Leclerc, 1989). This conservative bias makes sense if the result of the interpretation is the discounting of a region: if information is rendered inaccessible to higher-level processes (to at least some extent), this will be difficult to undo, and so should not be carried out if interpretation is at all doubtful.

The constraint against sudden changes in the luminance profile along the border (e) is similarly motivated: although it could result from a shadow falling on a dark region that just happens to have edges parallel to that of the shadow, it is more likely that the profile is part of the structure of the world, e.g. a line painted around the perimeter of the region. Interestingly, the subjective impression of a shadow is destroyed if lines inside are parallel to its border (Bühler, 1922); it may be that similar constraints are at play here.

The requirement that a shadowcaster be nearby (f) is another variation on this theme. The probability of a dark patch being a shadow increases with closeness to the caster, so that dark regions near a shadowcaster are most likely to be shadows cast by that object.

The constraint that the shadowcaster be a surface element (g) is due to simple considerations of physics: since light sources generally have some extension, they will give rise to penumbral effects that would simply wash out the thin shadow that a line element would generate. Note that the distinction between line and surface elements indicates that early vision is not exclusively concerned with surfaces, as has been sometimes proposed (e.g. He & Nakayama, 1992). Instead, it distinguishes between line and surface elements, with only the latter able to generate shadows.

Finally, the constraint that lighting be from above (l) can be justified by appeal to the regularities of the world. For the most part, lighting in the natural world is from above

(Lynch & Livingston, 1995), and a similar regularity also appears to exist in artificial environments—most lights are on ceilings rather than floors. It should be mentioned that this rule may instead reflect a constraint that the items be on a ground plane (i.e. a plane viewed from above) rather than a ceiling plane (i.e., a plane viewed from below). If so, a similar kind of explanation is involved, again appealing to regularities found in the world.[6]

Strategic Control

The discounting of regions identified as shadows accounts for the pattern of search slowdown encountered in these experiments. Interestingly, it may also account for another pattern encountered: in all 4 conditions where search slowdown was reliably encountered, the ratio of target-absent to target-present slopes in was high, averaging about 3.5. Indeed, in each of the 4 conditions, it was significantly higher than the ratio of 2.26 typical of most search tasks (Wolfe, 1998). These high ratios are consistent with a prolongation of search that compensates for the reduced accessibility of the discounted regions. As such, the higher-level strategic system that controls search also appears to be sensitive to the difficulties created by the discounting of cast shadows.

It is worth noting that high target-absent to target-present search ratios are found even for inverted cases, where search is quite rapid. Indeed, of the 4 conditions where search slowdown existed, search ratios were reliably higher for the inverted than for the upright cases ($p < .02$), with an average of 2.9 for upright and 4.0 for inverted. If shadows are indeed the basis for these high ratios, this finding suggests that there is a cost to the presence of shadows in these arrays even though they do not match the criteria for rapid analysis.

Other Systems

Although the studies here provide evidence for a rapid system concerned with shadow interpretation, other systems may exist as well, each with its own characteristics. Although the experiments here have emphasized the discounting of cast shadows, the information they contain can be used for a variety of purposes, such as establishing 3D surface shape or depth ordering between items (Baxandall, 1995; Cavanagh & Leclerc, 1989; Mamassian, Knill, & Kersten, 1998; Kriegman & Belhumeur, 2001). In addition, shadows have a clear phenomenological feel (Bühler, 1922; Hering, 1874/1964; Kardos, 1934). Experiments can be (and have been) based on each of these aspects of shadow perception, and it is not at all clear that a single system will necessarily emerge from such studies.

Already, some apparent contradictions have arisen between the results of various studies. For example, the relatively conservative constraints of the rapid system found here differ from those of the strategic system described above, and from those used for the extraction of 3D shape from shadow (Cavanagh & Leclerc, 1989). Part of this is likely due to the existence of different systems at different levels, with the operation of each tailored to the different computational resources available. But other factors are also involved. For example, shadows must be dark when used to extract 3D shape (Cavanagh & Leclerc, 1989), but can be white when used for motion in depth (Kersten, Mamassian, & Knill, 1997). Evidently, a variety of systems are being used, each likely specialized for a particular task.

It is also worth pointing out that the experiments here were based on the background being interpreted as a ground plane; a somewhat different set of constraints may apply to shadows cast on to a frontoparallel plane. Even more generally, cast shadows are only one kind of shadow: Other kinds exist as well, each with somewhat different constraints on its formation (see e.g. Baxandall, 1995; Cavanagh & Leclerc, 1989). Thus, there are many ways that the information in shadows can be used, and many constraints on their existence. A number of shadow systems could therefore exist at a variety of processing levels.

Conclusions

Much remains to be learned about shadow perception. But as the studies here have shown, a powerful way of studying shadows is to examine how they affect visual search. New methodological markers of shadows are important in the long run, for shadows are of interest not only in their own right, but also provide a way to improve our understanding of other mechanisms that could be used for other aspects of vision—e.g., the distinction between line and surface elements encountered here. To the extent that such approaches can be developed, new insights on human vision can emerge from out of the shadows.

Notes

1. The experimental protocol began by showing the observer a picture of upright stimuli (i.e., with the quadrilaterals attached to the bottoms of the rectangles, so that the quadrilaterals might be interpreted as shadows caused via lighting from above). Observers were then told that they would be looking for a shadow that "stuck out a little bit". For the cases where the inverted stimuli were tested, observers were told that in this condition they would see shadows that were upside down—i.e., the rectangles are attached to the ceiling and lit from below.

2. In almost all cases, the same pattern of results was obtained for *t*-tests based on *combined rates*, i.e, rates obtained by dividing the target-absent slopes of each observer in a particular case (upright or inverted) by the ratio of target-absent to target-present slopes for that case. The use of combined rates can partly compensate for the greater variability of target-absent responses (Wolfe, 1998) while still involving the data from these responses in the analyses. The only case where statistical significance was approached using the combined rates but not the target-present rates was Condition 2B.

3. Also called *obtruder* (e.g., Casati, 2000) and *casting object* (e.g. Mamassian et al., 1998).

4. Such shadows can result when the item casting the shadow is a transparent body with nonparallel surfaces. However, these are somewhat rare cases, being limited in the natural world to such things as water drops.

5. A line element will not cast a noticeable shadow if sufficiently thin: Virtually all lighting sources have some finite extent, leading to penumbral effects that will wash out a thin shadow. A line element with sufficient thickness would eventually be capable of casting a noticeable shadow, but this would be in virtue of its having surface structure. The lines used here are sufficiently thin that the shadows they would cast would likely be washed out.

6. Because the experiments described here rely on the assumption of a ground plane, this possibility cannot be directly tested here. One way of doing so might be via a set of experiments involving a different background structure, such as a frontoparallel plane. A complication here would be to ensure that a different set of constraints is not brought into play for such conditions.

References

Al-Haytham, I. (1000 / 1989). *The Optics of Ibn Al-Haytham. (Books 1-III).* A.I. Sabra, trans. London: The Warburg Institute.

Baxandall, M. (1995). *Shadows and Enlightenment* (pp. 1-15). New Haven, CT: Yale University Press.

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and Representation in Perception* (pp. 285-317). Hillsdale, NJ: Erlbaum.

Braje WL, Kersten D, Tarr MJ, & Troje NF. (1998). Illumination effects in face recognition. *Psychobiology*, **26**, 371-380.

Bühler K. 1922. Die Erscheinungsweisen der Farben. Jena: Fischer.

Casati R. (2000). The structure of shadows. In *Time and Motion of Socio-Economic Units*. (pp. 99-109). A Frank, J Raper, & JP Cheylan, eds. London: Taylor & Francis.

Cavanagh P. (1991). What's up in top-down processing? In *Representations of Vision: Trends and Tacit Assumptions in Vision Research*. (pp. 295-304) A. Gorea, ed. Cambridge, UK: Cambridge University Press.

Cavanagh P, & Leclerc YG. (1989). Shape from shadows. Journal of Experimental Psychology: Human Perception & Performance, **15**, 3-27.

Cunningham RK, Beck J, and Mingolla E. (1996). Visual search for a foreground object in continuous, naturalistic displays: The importance of shadows and occlusion. *Investigative Ophthalmology and Visual Science* , **37**, S299.

Elder JH., Trithart S, Pintilie G, & MacLean D. (1998). Rapid processing of cast and attached shadows, *Investigative Ophthalmology and Visual Science*, **39**, S853.

Enns JT, & Rensink RA. (1990). Influence of scene-based properties on visual search. *Science*, **247**, 721-723.

He ZJ, & Nakayama K. (1992). Surfaces versus features in visual search. *Nature*, **359**, 231-233.

Helmholtz, H. von. (1867 / 1967). *Treatise on physiological optics (Vol. 3).* In J. P. C. Southall (Ed. and Trans.). NY: Dover.

Hering, E.. (1874 / 1964). *Outlines of a Theory of the Light Sense*. LM Hurvich & D Jameson, transl. Cambridge, MA.: Harvard University Press.

Julesz, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neuroscience*, **7**, 41-45.

Kleffner DA, & Ramachandran VS. (1992). On the perception of shape from shading. *Perception & Psychophysics*, **52**, 18-36.

Kardos L. (1934). Ding und Schatten: Eine experimentelle Untersuchung über die Grundlagen des Farbensehens. *Zeitschrift für Psychologie*, **23**.

Kersten D, Mamassian P, & Knill DC. Moving cast shadows induce apparent motion in depth. *Perception*, **26**, 171-192.

Kriegman DJ, & Belhumeur PN. (2001). What shadows reveal about object structure. *Journal of the Optical Society of America A*, **18**, 1804-1813.

Lynch DK, & Livingston W. (1995). *Color and Light in Nature* (pp. 1-4). Cambridge: Cambridge University Press.

Mamassian P, Knill DC, & Kersten D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences*, **2**, 288-295.

Mitsudo H. (2002). Information regarding structure and lightness based on phenomenal transparency influences the efficiency of visual search. *Perception*, **32**, 53-66.

Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice Hall

Ramachandran, V.S. (1988). Perceiving shape from shading. *Scientific American*, **259**, 76-83.

Rensink RA, & Cavanagh P (1993). Processing of shadows at preattentive levels. *Investigative Ophthalmology & Visual Science*, **34**, S1288.

Rensink RA, and Enns JT (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, **102**, 101-130.

Rensink RA, and Enns JT (1998). Early completion of occluded objects. *Vision Research*, **38**, 2489-2505.

Treisman A, & Gelade G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

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

Wolfe JM. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, **9**, 33-39.

Figure Captions

Figure 1. Discounting of shadows. If the shadow is not discounted, irrelevant edges caused by its border can interfere with recognition. Discounting can largely reduce the effect of these irrelevant edges.

Figure 2. Example of a search display, using the standard stimuli (i.e., the stimuli of Experiment 1). (a) Upright case. The dark gray quadrilaterals attached to the vertically-oriented rectangles correspond to shadows cast by vertical posts onto the ground plane. (b) Inverted case. Quadrilaterals correspond to “upside-down” shadows on a ceiling plane.

Figure 3. Results of Condition 1A (i.e., the standard stimuli). Search for the unique orientation of the quadrilateral in the upright case is slower than in the inverted case. Error bars indicate standard errors of the mean. T = Target; D = Distractor.

Figure 4. Stimuli and results for the conditions of Experiment 1. Target is defined by a difference in the orientation of the quadrilateral attached to the rectangle. Numbers with dark gray backgrounds indicate a reliable difference in speeds between upright and inverted cases.

Figure 5. Stimuli and results for the conditions of Experiment 2. Target is defined by a conjunction of properties contained in the distractors. Numbers with dark gray backgrounds

indicate a reliable slowdown in the upright case. Numbers with light gray backgrounds indicate a reliable speedup in the upright case.

Figure 6. Stimuli and results for the conditions of Experiment 3. Target is defined by a difference in the orientation of the quadrilateral attached to the rectangle. Numbers with dark gray backgrounds indicate a reliable difference in speeds between upright and inverted cases.

Figure 7. Stimuli and results for the conditions of Experiment 4. Target is defined by a difference in the orientation of the quadrilateral attached to the rectangle. Numbers with dark gray backgrounds indicate a reliable difference in speeds between upright and inverted cases.

Figure 8. Stimuli and results for the conditions of Experiment 5. Target is defined by a difference in the orientation of the quadrilateral attached to the rectangle. Numbers with dark gray backgrounds indicate a reliable difference in speeds between upright and inverted cases.

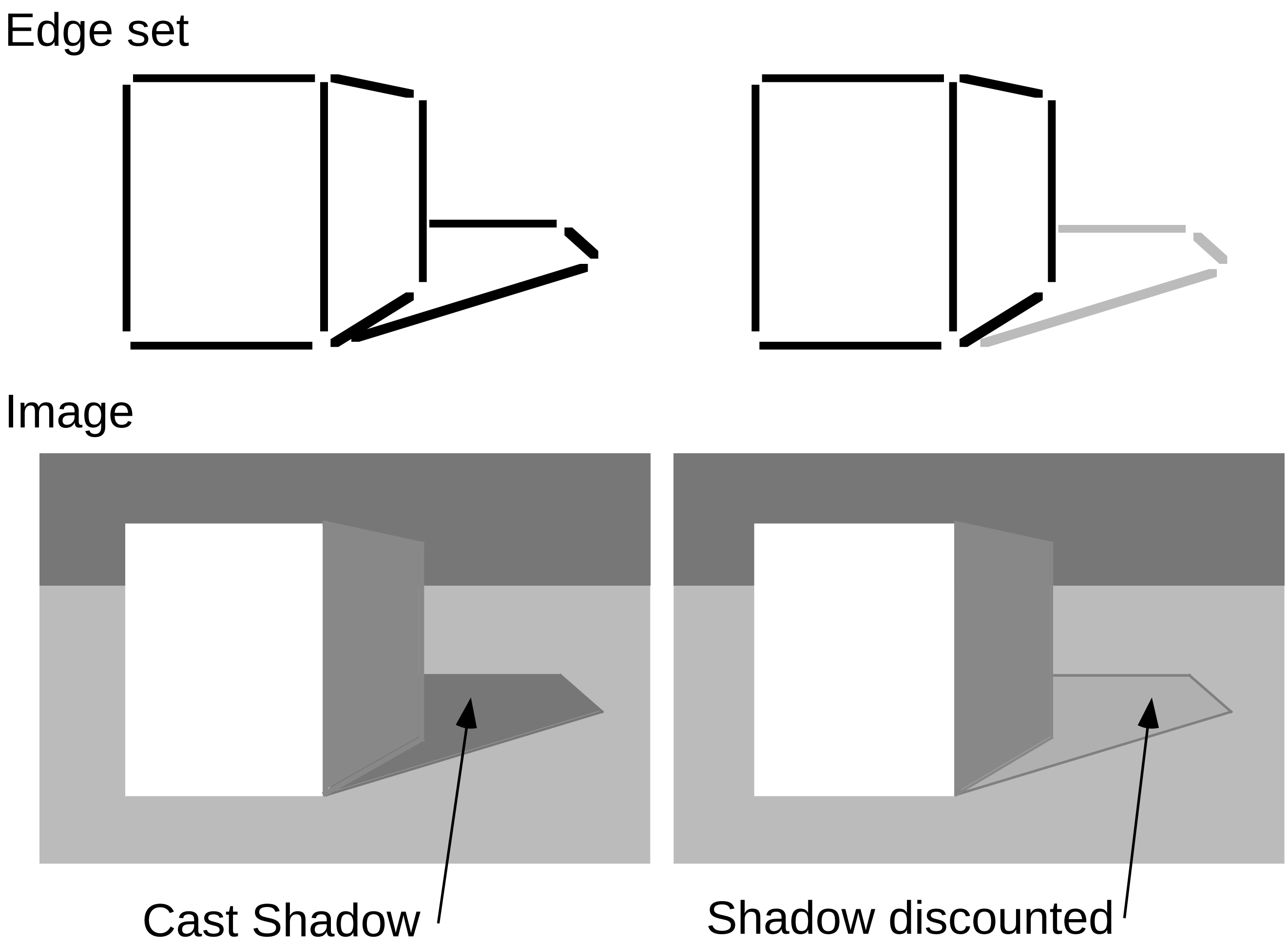
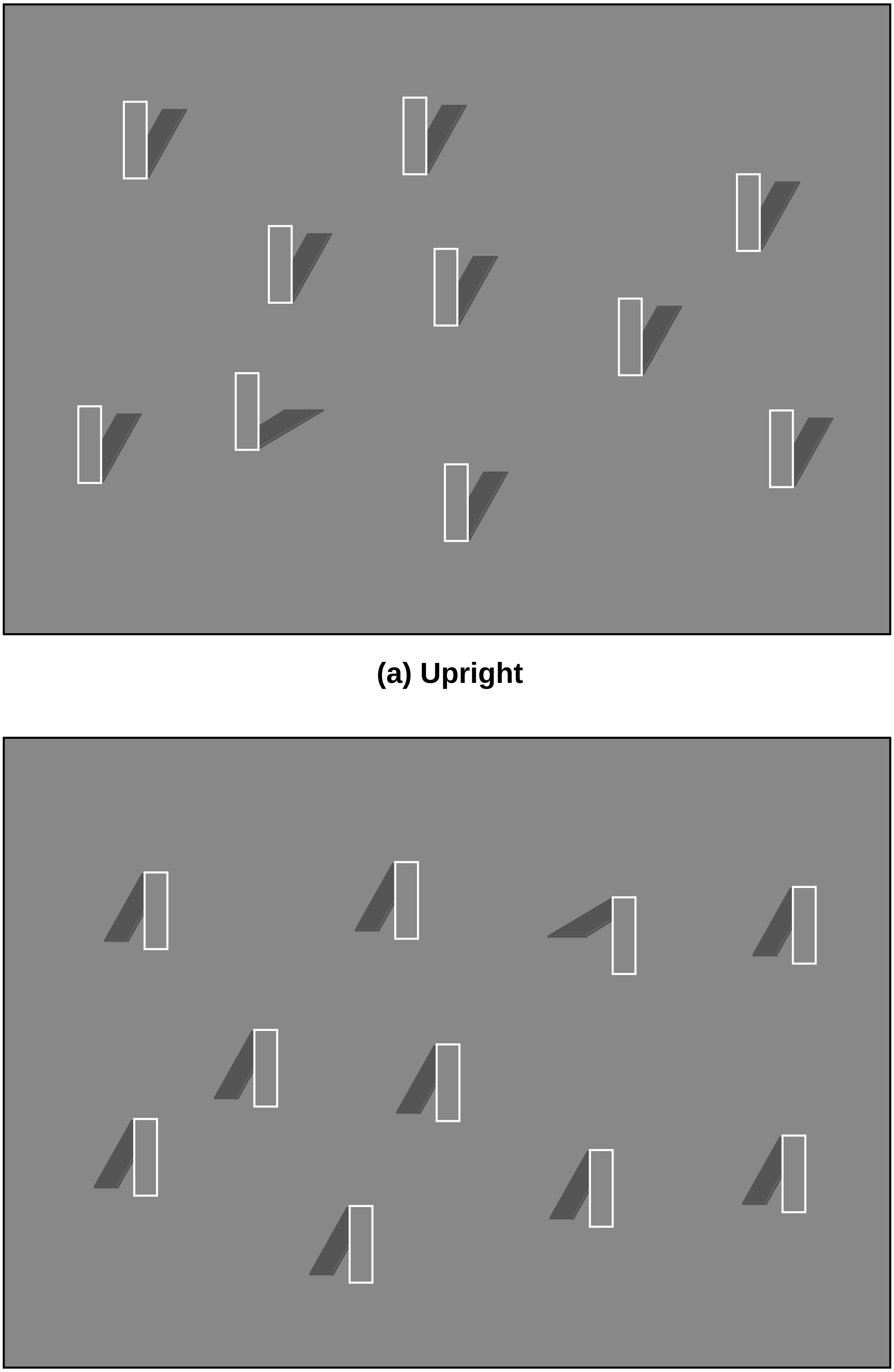


Figure 1



(a) Upright

(b) Inverted

Figure 2

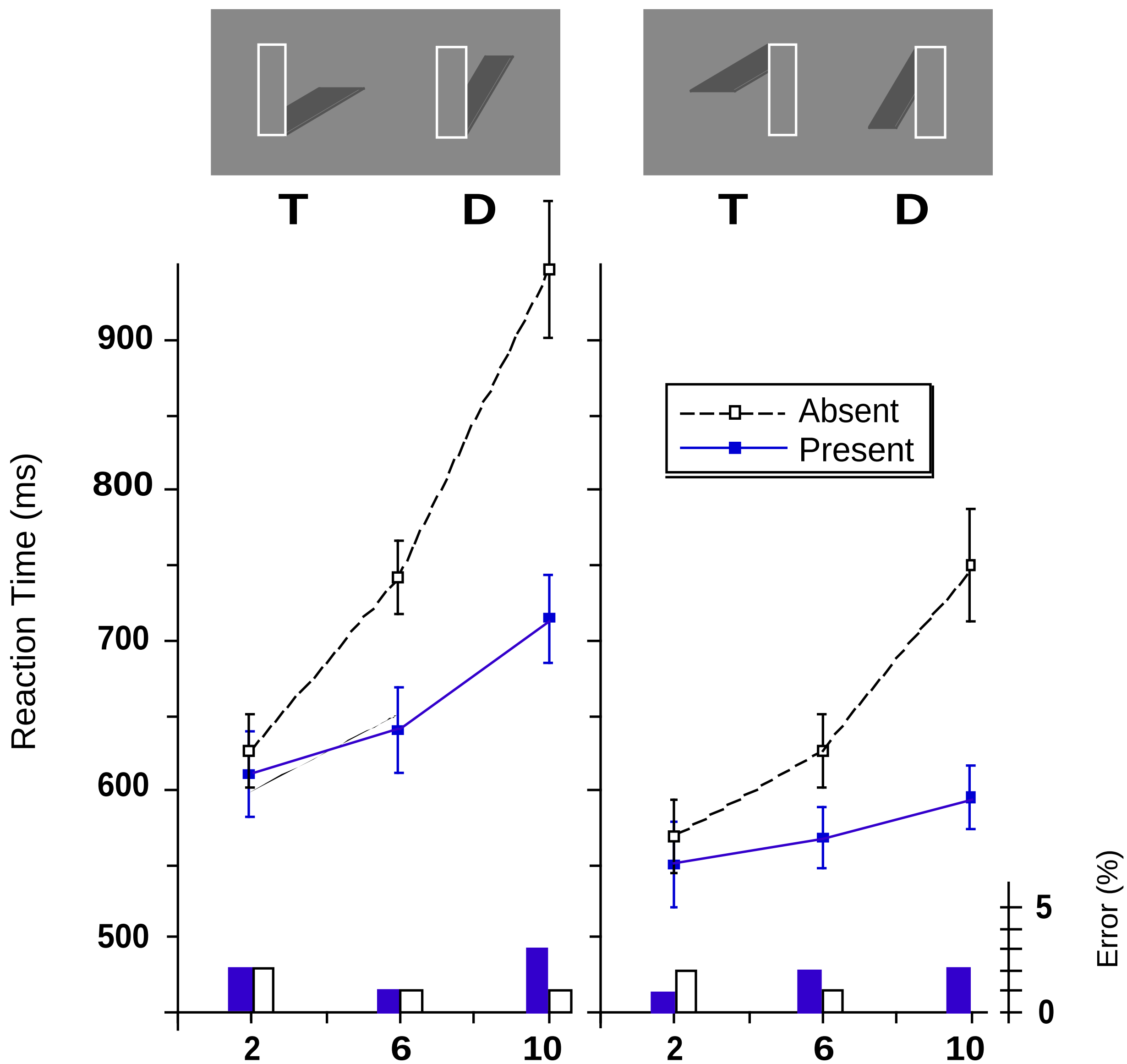

T
D
T
D
Reaction Time (ms)
900
800
700
600
500
Absent
Present
Error (%)
5
0
2
6
10
2
6
10

Figure 3

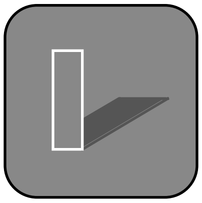
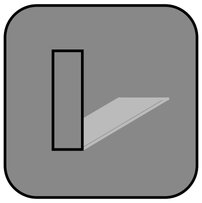
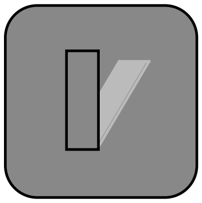
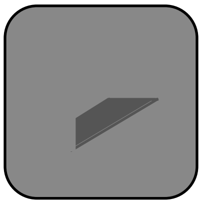
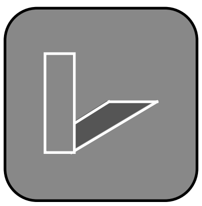
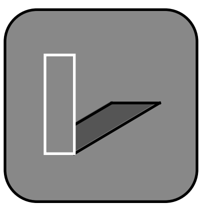
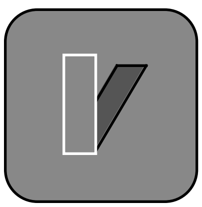
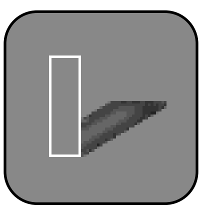
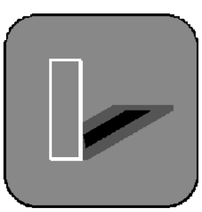
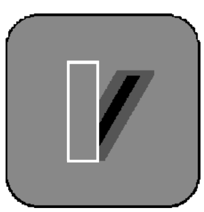
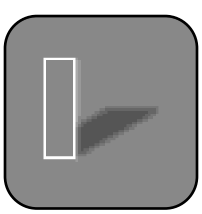
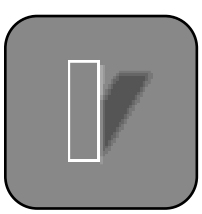
Cond	Target	Distractor	Up	Inv	*p*
1A			13.4 40.3	*4.8* *22.4*	.003
1B			5.1 12.1	*5.1* *11.4*	.99
1C			1.9 7.7	*1.4* *11.5*	.85
1D			5.5 11.2	*4.7* *10.2*	.65
1E			3.7 10.8	*2.4* *7.5*	.55
1F			2.4 12.6	*2.1* *10.0*	.80
1G			3.2 15.4	*3.0* *17.4*	.95
1H			14.3 42.9	*5.5* *22.0*	.003

Figure 4

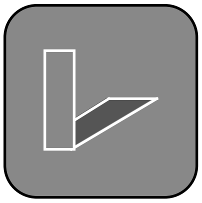
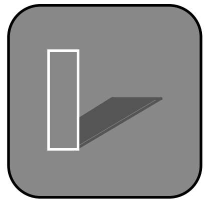
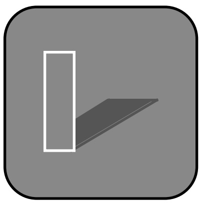
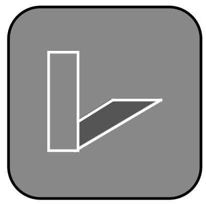
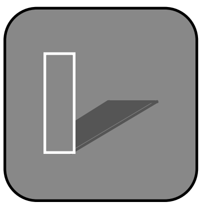
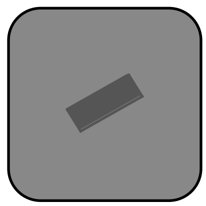
Cond	Target	Distrac 1	Distrac 2	Up	Inv	*p*
2A				30.4 68.2	*41.4* *96.0*	.03
2B				54.5 109.3	*47.3* *87.1*	.15
2C				32.5 53.1	*20.6* *38.2*	.015

Figure 5

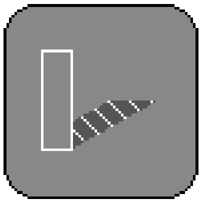
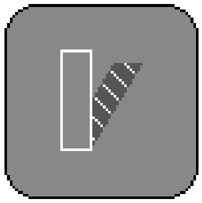
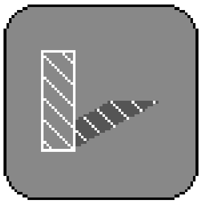
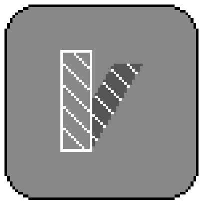
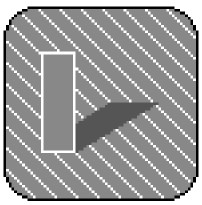
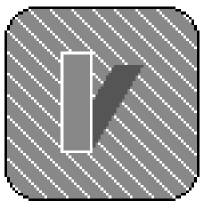
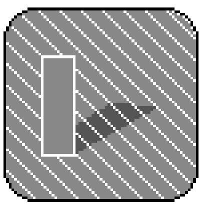
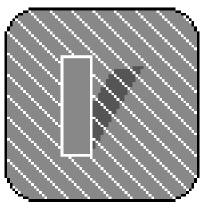
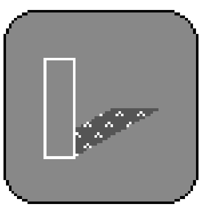
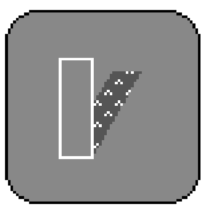
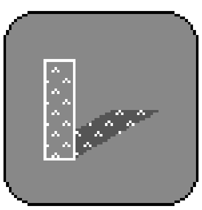
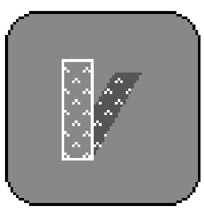
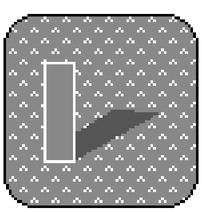
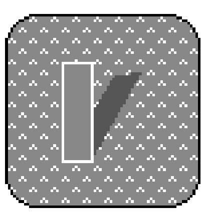
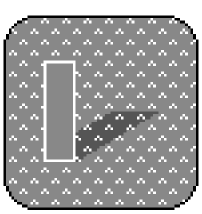
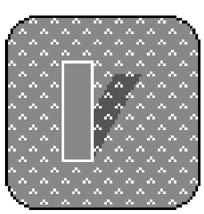
Cond	Target	Distractor	Up	Inv	*p*
3A			4.9 18.6	*9.2* *19.3*	.20
3B			4.1 23.9	*5.6* *20.2*	.50
3C			4.3 16.6	*2.6* *9.5*	.25
3D			12.5 34.9	*4.4* *18.1*	.025
3E			13.9 39.6	*12.0* *35.9*	.25
3F			7.4 19.9	*7.2* *14.5*	.95
3G			3.5 19.6	*4.4* *13.3*	.60
3H			12.0 34.3	*6.5* *22.2*	.025

Figure 6

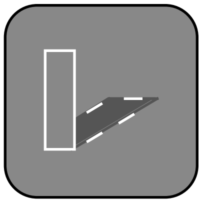
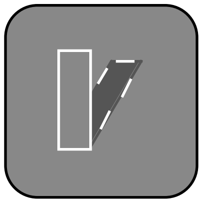
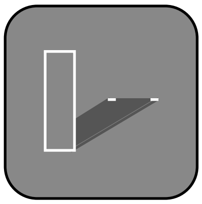
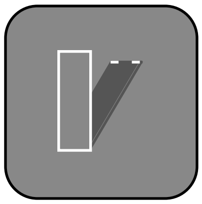
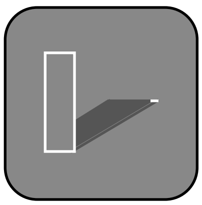
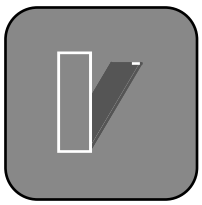
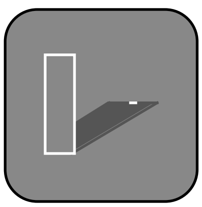
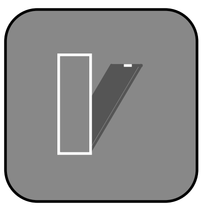
Cond	Target	Distractor	Up	Inv	*p*
4A			6.9 17.2	*7.7* *18.8*	.75
4B			3.2 17.2	*3.0* *17.1*	.95
4C			6.5 18.7	*6.1* *11.8*	.85
4D			4.1 22.3	*1.5* *17.1*	.30

Figure 7

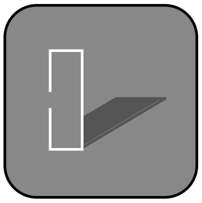
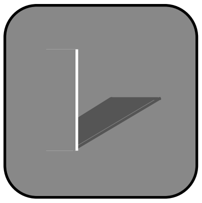
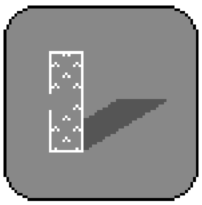
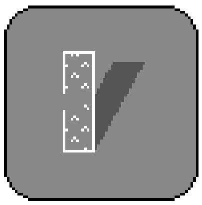
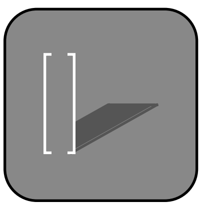
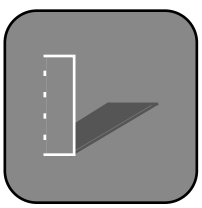
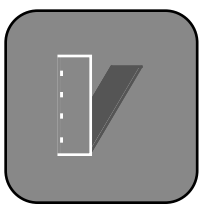
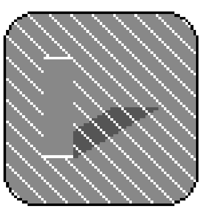
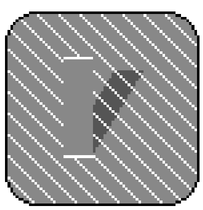
Cond	Target	Distractor	Up	Inv	*p*
5A			6.0 18.4	*5.1* *16.1*	.60
5B			1.9 14.7	*3.7* *9.6*	.50
5C			8.8 33.8	*4.0* *17.0*	.20
5D			3.3 22.7	*5.7* *15.2*	.30
5E			3.8 20.7	*3.1* *12.4*	.80
5F			6.1 47.2	*4.5* *22.5*	.55

Figure 8