What makes cast shadows hard to see?

Gillian Porter
Department of Experimental Psychology, University of Bristol, Bristol, UK

Andrea Tales
Department of Experimental Psychology, University of Bristol, Bristol, UK

Ute Leonards
Department of Experimental Psychology, University of Bristol, Bristol, UK

Visual search is slowed for cast shadows lit from above, as compared to the same search items inverted and so not interpreted as shadows (R. A. Rensink & P. Cavanagh, 2004). The underlying mechanisms for such impaired shadow processing are still not understood. Here we investigated the processing levels at which this shadow-related slowing might operate, by examining its interaction with a range of different phenomena including eye movements, perceptual learning, and stimulus presentation context. The data demonstrated that the shadow mechanism affects the number of saccades during the search rather than the duration until first saccade onset and can be overridden by prolonged training, which then transfers from one type of shadow stimulus to another. Shadow-related slowing did not differ for peripheral and central search items but was reduced when participants searched unilateral displays as compared to bilateral ones. Together our findings suggest that difficulties with perceiving shadows are due to visual processes linked to object recognition, rather than to shadow-specific identification and suppression mechanisms in low-level sensory visual areas. Findings are discussed in the context of the need for the visual system to distinguish between illumination and material.

Keywords: shadows, light from above, visual search, eye movements, perceptual learning, context


Introduction

In a typical scene, objects will occlude light produced from a specific directional light source, resulting in regions of shading on the surfaces of the objects themselves (“attached shadows”) and shadows cast upon other surfaces (“cast shadows”). Both types of shadow can provide valuable information helping the viewer to interpret what they see. In particular, attached shadows provide salient information to help in understanding an object’s 3D shape (e.g., Attwood, Harris, & Sullivan, 2001; Ramachandran, 1988; Sun & Perona, 1996), with ambiguity often resolved by assuming a single light source above the head (Kleffner & Ramachandran, 1992; Ramachandran, 1988). Cast shadows have been shown to have value for understanding spatial layout of a scene (Allen, 1999; Hubona, Shirah, & Jennings, 2004), especially when it involves motion (Kersten, Mamassian, & Knill, 1997; Mamassian, Knill, & Kersten, 1998). However, cast shadows also provide “noise”, which might potentially impair understanding of the visual input, and indeed other studies suggest that shadows can be difficult to process; for example, incongruent shadows often go unnoticed in pictorial scenes and art (Jacobson & Werner, 2004; Ostrovsky, Cavanagh, & Sinha, 2005). The present paper further explores the apparent problems in processing cast shadows under certain circumstances (Rensink & Cavanagh, 2004), with a particular focus on the possible underlying mechanisms.

Making use of the implicit light-from-above assumption (Adams, 2007; Ramachandran, 1988), Rensink and Cavanagh (2004) demonstrated that visual search for items that could be perceived as shadows was slower than search for equivalent items that were not shadow-like. Specifically, when search items could be interpreted as rectangular posts with cast shadows behind and the target “shadow” was of a slightly discrepant angle (see Figure 1a), search slopes were steeper than for displays using the same stimuli inverted, in which the “shadow” no longer complied with light-from-above assumptions and was not therefore interpreted as shadow-like. An upright–inverted difference was recorded only if the “shadow” region possessed specific qualities, and Rensink and Cavanagh identified many constraints necessary for a percept to be interpreted as a shadow (e.g., the candidate region must be darker than the background and have no hard outline). They concluded that cast shadows are rapidly identified and then (at least partially) discounted to minimize noise in the visual scene. As a consequence, it is more difficult to conduct perceptual judgments on shadowy as compared to non-shadowy image regions.
Other visual search studies for shadows have confirmed this effect. Elder, Trithart, Pintilie, and MacLean (2004, Experiment 3) used circular stimuli, each with an adjoining crescent-shaped "cast shadow", which could be darker (shadow-like) or lighter than the background (control). Participants searched for a discrepant target crescent, oriented at 180° to distractors. Although Elder et al. primarily discussed search speeds for darker compared with lighter shadows, more relevant to this paper was their finding of a search asymmetry for upright versus inverted items specific to darker shadows, analogous to Treisman and Souther's (1985) observation that it is more difficult to find an O among Qs than a Q among Os. In Elder et al.'s study, search for a target with the shadow top left (i.e., inverted, according to light-from-above constraints) among distractors with the shadow bottom right (i.e., upright) led to flatter search slopes than search for an upright target among inverted distractors. That is, in the context of Rensink and Cavanagh's (2004) claims, an (easily perceived) inverted shadow "popped-out" among (hard-to-perceive) upright ones, whereas a (hard-to-perceive) upright shadow was well concealed among easily perceived inverted distractors (see also Rensink & Cavanagh, 1993). Similarly, Lovell, Gilchrist, Tolhurst, and Troscianko (2009) demonstrated faster search of natural images of heterogeneous pebbles with upright (light-from-above) shadows for one with a discrepant shadow inverted 180° (light-from-below) than vice versa. Lovell et al. also showed how the orientation difference between target and distractors affects search speeds. With smaller orientation differences (120°–60°), the target no longer appeared light-from-below and could be found equally fast among light-from-above distractors to when the displays were inverted. Once the orientation difference became small (30°), however, search was slower with the upright (light-from-above for both target and distractors) than the inverted displays, just as reported by Rensink and Cavanagh (2004). This is in keeping with shadow processing being coarsely scaled, so that only large shadow discrepancies can be readily perceived (Lovell et al., 2009; Mamassian, 2004).

Figure 1. Stimuli. (a) Upright, shadow-behind, target (left) and distractor (all experiments). (b) Inverted, non-shadow, target (left) and distractor (Experiments 1 and 2). (c) Upright, shadow-in-front, target (left) and distractor (Experiments 2 and 5). (d) Whole display, light-from-above-right lighting context, congruent shadows (Experiment 5).

While all of these reports (Elder et al., 2004; Lovell et al., 2009; Rensink & Cavanagh, 2004) agree that the visual system has more difficulty in inspecting a shadow-like item than the exact same visual input inverted, the underlying mechanisms and processing levels remain unclear. Rensink and Cavanagh (2004) suggested that, since shadow-like status can slow down otherwise very fast search, shadow identification and discounting involves rapid, parallel mechanisms and thus probably early levels of vision. To retrieve discounted shadow information, attention is then needed, explaining steeper search slopes. Nevertheless, Rensink and Cavanagh point out that higher level knowledge about the constraints that define a shadow must also be involved, so relevant information must circulate between levels (see, e.g., Di Lollo, Enns, & Rensink, 2000). Elder et al. (2004) proposed that shadows are not discounted but contribute to 3D object percepts, and it is these that can be rapidly processed, rather than the shadows themselves. This fits with the idea that rapidly formed grouped structures (e.g., objects) can be readily accessible, while their component parts are not.
(Rensink, 2000; Rensink & Enns, 1995). Lovell et al. (2009) also questioned Rensink and Cavanagh’s assertion that shadows are discounted, given the rapidity with which some shadow-based discriminations can be made. Instead they advocate a very early shadow identification filter, applied to all types of scenes, allowing “quick and dirty” coarse-scaled shadow-related judgments, but requiring other mechanisms to then be used if the shadow system fails (e.g., to identify small shadow discrepancies).

In a series of behavioral experiments, all based upon Rensink and Cavanagh’s (2004) shadow-search task, we sought to further investigate the mechanisms underlying shadow perception and the processing levels at which these mechanisms might operate. In particular, we tried to narrow down the location of Rensink and Cavanagh’s (2004) early system for discounting shadows by behaviorally testing whether this might involve processing in early sensory visual areas. We noted that Rensink and Cavanagh (2004) focused mainly upon search slopes (i.e., search speed per item) to distinguish between shadow and non-shadow search, yet there also appeared to be a global effect of shadow presence upon search speeds, independent of set size. Thus, the search function’s mean y-value at $x = 0$ (the intercept) or at the mean x-value (the centercept, see Wainer, 2000) showed similar shadow-related differences to slopes (Elder et al., 2004; Lovell et al., 2009; Rensink & Cavanagh, 2004). Longer intercepts or centercepts when distinguishing small shadow discrepancies compared with non-shadows might support Lovell et al.’s (2009) idea of an early, distinct, coarse-scaled shadow-processing stage acting and failing before search mechanisms apply. Taking this as a starting point, we first used eye movement recordings to investigate such a possibility. Next, we looked at whether the specific disadvantage for accessing shadow-like items, as seen in longer processing times, can be overcome with prolonged practice, and whether such perceptual learning would transfer to other testing conditions, thus most likely excluding the primary visual cortex as shadow processor. To investigate the role of low-level visual areas in shadow processing, we then looked at whether the elimination of eye movements affects shadow-related slowing, and whether it varies with bilateral or unilateral visual input. We further explored the potential interaction of shadow accessibility with object-based attentional biases. Finally, we manipulated the explicit lighting context to test whether light-from-above assumptions could be overruled.

### Experiment 1: Eye movements

Instead of a shadow discounting mechanism, Lovell et al. (2009) proposed a model involving a very early shadow identification stage for all types of stimuli. If shadows were present, but the coarse resultant information from such a processing stage was inadequate for the task at hand (e.g., for discriminating similar shadows), then the processing mode would have to switch to the more fine-grained analysis process of non-shadow input. As suggested by Lovell et al., this switch would cost valuable processing time. If such an early shadow identification stage exists, we would expect a global (non-search) related slowing effect upon upright shadow RTs relative to inverted in Rensink and Cavanagh’s (2004) paradigm; indeed, Rensink and Cavanagh’s prolonged search intercepts for shadows might suggest exactly this. However, prolonged search intercepts could equally result from other, later mechanisms for making a decision related to the shadow-like input.

A dependent measure more specific to early sensory processing rather than to decision making is the first fixation duration (i.e., the length of time from stimulus onset to the first saccade). First fixation duration is thought to reflect the complexity of a stimulus display: the less complex, the shorter the first fixation (Leonards & Scott-Samuel, 2005; Zelinsky, 2001; Zelinsky & Sheinberg, 1997).

The first experiment therefore explored the eye movement patterns accompanying upright and inverted shadow search in comparison with a non-shadow control. We used stimuli with small orientation discrepancies based upon those used by Rensink and Cavanagh (2004), so some displays involved a target and distractors that were all shadow-like, while the other displays had only items that were not shadow-like. If the failure of an early shadow-processing system contributes to the slowing recorded by Rensink and Cavanagh (2004), then alongside the RT differences we would expect to find longer initial fixation durations in upright than inverted shadow conditions but no upright–inverted difference for the white non-shadow control condition.

### Methodology

#### Participants

Twelve psychology students (three males) aged 18–20 (mean 19.2) years participated in exchange for course credits. All gave their written consent and procedures were approved by the Faculty of Science Ethics Committee, University of Bristol.

#### Stimuli and tasks

Participants were asked to perform a standard target-present/target-absent visual search task. They searched for an odd-one-out within a display and pressed one key (the left arrow) if they found one, and another key (right arrow) if they decided all items were identical. They were instructed to press the appropriate response key immediately when the decision was made. Eye movements were recorded throughout, along with reaction times and accuracy. A tone indicated an incorrect response.
Stimulus elements consisted of a rectangular shadow caster, 17 mm tall by 5.5 mm wide, with an oblique “shadow” attached. Distractor elements were all identical, with the “shadow” oriented at 30° to the vertical, while the target element (one per target-present display) had the “shadow” oriented at 60° to the vertical. The area of the “shadow” was equated across target and distractor elements (see Figure 1). Displays involved set sizes of 4, 8, or 12 elements, located within a virtual 5 x 5 cell grid in which location centers formed a 10.2 cm square. Item density was approximately equated across set sizes by having items for set sizes 4 and 8 located within a virtual 3 x 3 or 4 x 4 sub-grid, respectively, within the main grid. Randomized jitter of up to ±6 mm in both horizontal and vertical directions was added to item locations to avoid an appearance of regularity. Using Matlab, stimulus arrangements were pre-prepared to ensure distributions with no touching elements and to allow target location to be evenly distributed across the 5 x 5 grid for each set size and each task.

Two stimulus types were included, each with upright and inverted versions: shadow, a white outlined shadow caster and dark shadow (Figure 1a), and white non-shadow, a black outlined shadow caster and light shadow (Figure 1b). Pre-testing of alternative stimuli showed no effect of having, for example, an outlined rather than a solid shadow caster. Background luminance was measured as 43.2 cd/m², with the dark shadow recording a Michelson contrast of 20% relative to this. The white non-shadow was of slightly lower contrast, at 15% relative to the background.

Procedure

Participants were tested in a single test session in which they completed four blocks of each of the two tasks, separated by rest periods. Displays were presented using Matlab and the Psychophysics Toolbox (Brainard, 1997) on an 18" LCD monitor viewed at 57 cm, with head position maintained using a chinrest, in a dimly lit room. Following an initial practice block of 48 trials, an Eyelink II headset (S.R. Research) was fitted and adjusted to give clear readings. Before each of the eight test blocks, a calibration and validation procedure for eye tracking was completed using a 9-point calibration grid. Each stimulus display was preceded by a uniform display with a central fixation cross, with fixation necessary before the trial could commence.

Two sets of display arrangements were used, matched in terms of item distributions across different trial types (target presence and orientation). Stimulus sets were counterbalanced across the two tasks and participants so that, across the sample, identical stimulus arrangements were used for each task and for each stimulus orientation. Each block comprised 4 initial practice trials, not included in analysis, then 5 trials of each combination of stimulus orientation (upright or inverted), set size, and target presence, giving 60 trials proper, presented in randomized order. Point-of-gaze coordinates were recorded at a frequency of 500 Hz and with an accuracy of <0.3°, and saccades identified from coordinate changes representing a velocity of 30°/s and acceleration of 8000°/s² or more. In total, each participant completed 20 target-present and 20 target-absent trials of each of three set sizes for each of two stimulus orientations for each of two tasks, giving 480 trials.

Analysis

Median correct reaction times (RTs) were calculated for each individual for each combination of condition and set size. Following Rensink and Cavanagh (2004), these were first used to calculate search slopes that were compared for upright versus inverted versions of each stimulus type using paired t-tests. The eye movement variables, specifically first fixation duration, were likewise calculated for each individual for each condition and set size, again based upon correctly answered trials only. First fixation analyses also excluded any anticipatory trials (first fixation duration <80 ms). The data were entered into 4-way (stimulus type x orientation x target presence x set size) ANOVAs, looking specifically for interactions of stimulus type (shadow or non-shadow) with stimulus orientation.

Results and discussion

The RT data replicate those of Rensink and Cavanagh (2004) in showing a behavioral disadvantage for search among upright shadow elements: Target-present search slopes were significantly steeper for upright than inverted displays in the shadow condition (7.16 ms/item vs. 0.446 ms/item; t(11) = 3.78, p = 0.003) but not for the white non-shadow displays (upright, 1.38 ms/item vs. inverted, 3.57 ms/item; t(11) = −1.05, p = 0.318). The same was true for averaged target-present/absent slopes (shadow, t(11) = 4.02, p = 0.002 and white non-shadow, t(11) = −1.88, p = 0.087).¹ Intercepts did not show significant differences; however, there was a tendency toward longer absolute RTs with upright than inverted shadow displays. Error rates were low (mean 2.1%) and did not differ significantly by condition. In all these respects, our participants performed similarly to those of Rensink and Cavanagh (2004).

We predicted that slower RTs to upright than inverted shadow-like stimuli would be reflected in longer first fixation durations for the upright shadow condition than the inverted or control conditions, if a very early shadow-processing system was involved (Lovell et al., 2009). As a basis for this comparison, we analyzed total RTs without separating intercepts and slopes. Using ANOVA, the crucial stimulus type x orientation interaction was significant for total RTs (see Figure 2a; F(1, 11) = 9.52, p = 0.010): RTs were slower to upright dark shadow
than upright white non-shadow items (755 ms vs. 679 ms, $p = 0.02$, Fisher’s LSD), while RTs to inverted items did not differ between conditions (721 ms and 697 ms, respectively).

For first fixation durations, a significant stimulus type × orientation interaction was also recorded ($F(1, 11) = 5.10, p = 0.045$), but the interesting comparisons were non-significant in post-hoc testing (dark shadow, upright vs. inverted, $p = 0.189$; upright dark shadow vs. upright white non-shadow, $p = 0.459$, Fisher’s LSD). Moreover, the pattern was the opposite of that required to explain RTs (see Figure 2b). The eye movement data therefore fail to provide supportive evidence for there being an extra shadow-recognition stage very early in image processing (Lovell et al., 2009). Note, however, that first fixation duration varied markedly with set size ($F(2, 22) = 74.8, p < 0.001$; 202 ms, 233 ms, and 264 ms for sets of 4, 8, and 12 items, respectively, all differences, $p < 0.05$, Fisher’s LSD), so the lack of a sizeable difference by stimulus type and orientation cannot be attributed to poor sensitivity of this measure to sensory aspects of image complexity.

Although first fixation durations failed to reflect the longer RTs for upright shadows, we reasoned that the observed shadow-specific slowing must emerge elsewhere in the eye movement patterns, either as a lengthening of later fixation durations or as an increase in the number of saccades for upright shadows (see Hooge & Erkelens, 1996, 1999; Zelinsky & Sheinberg, 1997), or both. Further analysis showed that the duration of later fixations excluding the first was, like first fixations, invariant to stimulus type × orientation manipulations (interaction, $p = 0.773$). This suggests that there is no slowing associated specifically with response preparation processes in the shadow-like condition, nor with general processing of local shadow-like regions. In contrast, we found a significant stimulus type × orientation interaction for the total number of saccades made per trial ($F(1, 11) = 5.99, p = 0.032$): more saccades were made in response to upright shadow than upright non-shadow displays (2.85 vs. 2.50; $p = 0.028$, Fisher’s LSD, Figure 2c). Since number of saccades is generally taken to reflect voluntary, strategic factors (Zelinsky & Sheinberg, 1997), this suggests that accessing the orientation of shadows requires a change in search strategy. This is in line with Rensink and Cavanagh’s (2004) observations that upright shadow search seemed to involve a different strategy from inverted, as indicated by surprisingly high target-absent/target-present search slope ratios.

What form might this change in strategy take? One possibility is that it represents a strategic need to recheck the target before reaching a decision, specific to upright
shadow displays. However, although there were slightly fewer single-fixation trials for upright shadow-like (mean 15.8%) than inverted (17.2%) or white upright (19.4%) control trials, this was not significant (interaction, p = 0.280). Moreover, point-of-gaze analysis showed that the number of saccades per target-present trial landing within the target region did not vary by stimulus type or orientation (interaction, p = 0.471). These analyses suggest that the shadow-like condition is associated either with more checking of the target, or conversely with a reduced tendency to fixate shadow-like regions, as might occur through inhibition mechanisms. The key stimulus type × orientation interaction did, however, emerge in the percentage of trials involving four or more eye movements (F(1, 11) = 12.0, p = 0.005), with more trials involving multiple eye movements for upright shadow than upright non-shadow displays (24.6% vs. 15.7%; p = 0.025, Fisher’s LSD). This appears to represent a general rather than occasional increase in saccades during upright shadow search, given that the variability (standard deviation) of saccade numbers was not significantly greater for this condition (interaction, p = 0.100). We conclude here that search in the upright shadow condition is characterized by trials requiring a larger number of saccades, but the cause of this prolongation of the search process remains unclear. We will return to the role of eye movements in Experiment 3.

### Experiment 2: Perceptual learning

Eye movement analysis in Experiment 1 failed to provide support for a shadow-specific, very early stage of processing (Lovell et al., 2009), globally separating shadows from non-shadows in visual scenes before actual search processes come into play. Instead, Experiment 1 indicated that the slowing associated with searching among light-from-above shadows arose from an increase in the number of saccades on shadow-like trials relative to non-shadow-like trials. With number of saccades thought to be under strategic (and attentional) control (Zelinsky & Sheinberg, 1997), we reasoned that it might be possible with prolonged practice to overcome the processes that discriminate shadow search from non-shadow search.

Performance of many types of perceptual tasks improves with training. Moreover, practice may change the basic relationships between test conditions—for example, some visual search tasks with initially steep search slopes can, after training, be performed equally fast regardless of set size; in other words, they become efficient or “parallelize” (Leonards, Rettenbach, & Sireteanu, 1998; Sireteanu & Rettenbach, 1995, 2000). In contrast, other task types maintain a set size effect even after extensive practice (Leonards, Rettenbach, Nase, & Sireteanu, 2002; Sireteanu & Rettenbach, 2000). Parallelization is thought to rely mostly on changes in general search strategies and task-related spatial allocation of attention (Ahissar & Hochstein, 2000; Sireteanu & Rettenbach, 2000) rather than on changes of sensory discrimination thresholds. Therefore, in Experiment 2 we tested the hypothesis that the disadvantage for searching upright shadows relative to inverted/non-shadow control conditions can be overcome with prolonged experience. Not only set size effects but more importantly the orientation × stimulus type interaction seen in previous experiments should disappear over many testing blocks.

If the mechanism by which shadows are suppressed can be overcome with practice, we could then start to narrow down its “location” in the visual system by investigating whether such learning would transfer to untrained stimuli. Learning of low-level sensory discriminations, for example, transfers only to very similar test conditions (Ahissar & Hochstein, 1993), whereas learning that transfers to tasks using quite different stimuli tends to indicate that higher level processes have been trained (Ellison & Walsh, 1998; Sireteanu & Rettenbach, 1995, 2000). We predicted that any shadow-related learning would transfer fully across stimuli, indicating the involvement of cortical processes beyond the primary visual cortex (V1).

### Methodology

Two of the authors (GP and UL) and one naive subject (AW) participated. On twelve testing sessions, each separated by at least half a day, each participant completed target-present/target-absent visual search tasks with four different stimulus types. These four stimulus types, all with upright and inverted variants, included the white non-shadow stimuli of Experiment 1 as a control and three shadow conditions varying in contrast. Specifically, a lower and a higher contrast shadow condition (11% and 45% contrast between shadow and background) were included alongside the 20% contrast shadow condition of Experiment 1. This both increased participants’ experience of the shadow discrimination task for more concentrated learning and allowed an investigation of how shadow contrast affects the disadvantage for upright shadows. Testing was conducted on a Toshiba Tecra M4 laptop computer, with stimulus dimensions reduced to approximately 80% of those in Experiment 1. Luminances were consistent with those used in Experiment 1. In each testing session, the participant completed two blocks of each of the four tasks in a counterbalanced order, which was rotated across testing sessions. After 12 sessions, this totaled 11,520 trials or 1,440 of each orientation of each stimulus type. Four matched sets of display arrangements were used, counterbalanced across tasks and sessions, and rotated across participants.

To examine whether learning transferred to different angled stimuli, GP and UL then participated in four further sessions in which stimuli were changed so that the...
“shadow” was a parallelogram positioned below the shadow caster (i.e., apparently on the ground plane in front of it, at 150° for distractors/120° for targets; see Figure 1c). Only 20% contrast shadow and white non-shadow control conditions were used. Perpendicular width and the visible area of the “shadow” were closely matched across all items in order to minimize obvious visual differences in shadow luminance, width, or length between the original displays and the transfer ones. As before, each session involved two blocks with each stimulus type, in counterbalanced order and with 60 trials per block.

Results and discussion

Accuracy for all participants in the initial learning phase was, as expected, very high (over 90% for every condition at every session) and did not change markedly with practice. All participants showed learning effects in terms of performance speed, particularly over the early sessions but apparently continuing right through to session 12 (see Figure 3, top panel).

For each individual, RTs for correct responses to individual display arrangements were entered into a 4-way mixed ANOVA for each session, using set size and stimulus type as independent measures and orientation and target presence as repeated measures. Again, the focus was the stimulus type × orientation interaction, which we expected to be readily apparent in the early sessions, with slower upright than inverted search for the shadow tasks but not the control task. We predicted that the upright–inverted difference for shadow-like stimulus conditions would disappear over time, if the shadow suppression mechanism was sensitive to perceptual learning. The lower panels of Figure 3 illustrate the pattern of results, shown as the (upright–inverted) RT difference for each stimulus type, for each participant separately.

In keeping with previous findings, RTs to shadow-like stimuli were consistently longer when upright than inverted, whereas RTs to control stimuli showed much less of an orientation difference. In addition, there was a tendency toward a more marked orientation effect for lower than higher contrast shadows, approximately in proportion to overall RTs (see Figure 3, first 12 sessions). Despite the inevitable noise in trial-by-trial data, the predicted orientation × stimulus type (20% contrast shadows versus white non-shadows) interaction reached significance on most of the first 12 sessions for UL and on nearly half of sessions for AW and GP, considerably more than would occur by chance. However, this interaction lost significance as search speeds became faster, and although the overall speed disadvantage for upright relative to inverted shadows was maintained throughout for all three participants (shadow-related bars remain positive across sessions 1–12 in Figure 3), this no longer distinguished shadows from white non-shadows by session 12. Thus, in line with our hypothesis, shadow-specific slowing disappeared with prolonged training for all three participants.

We had anticipated, based upon its manifestation in increased saccade numbers, that the shadow-specific upright–inverted difference could be overcome by increasing familiarity with and hence salience of the relevant shadow orientation difference, rather like the processing of some difficult feature discriminations (Leonards et al., 2002, 1998). Our data conform to this prediction. In particular, our data suggest that the “lighting-from-above” assumption (Adams, 2007; Ramachandran, 1988), which creates the interpretation as shadows in the first place, is accessible to perceptual training inasmuch as people can learn either to ignore it or to override its effects upon perception. As such, the data are somewhat in keeping with Adams, Graf, and Ernst’s (2004) data showing that people can learn to adjust the assumed light source position.

The rightmost sections of Figure 3 plot the RTs after stimulus change in the learning transfer part of this experiment. Error rates did not differ markedly before and after stimulus change (3.4% then 5.0% for GP and 0.8% then 0.7% for UL). Although absolute RTs slightly increased when the new discrimination was introduced (Figure 3, top panel), there was no tendency for the shadow-specific upright–inverted RT difference to return (Figure 3, lower panels). These data therefore indicate full transfer of the shadow-specific learning from one angle discrimination to another, suggesting that it occurs at cortical levels beyond V1 (Sireteanu & Rettenbach, 1995, 2000).

The occurrence of transferable learning is likely to reflect the more time-effective deployment of attentional resources toward shadows with practice (Leonards et al., 1998). This would be compatible with the idea that attention is needed to retrieve discarded shadow information (Rensink & Cavanagh, 2004), and if so, the exact shape of the shadow information should not matter, as fits with the transfer experiment. Alternatively, learning could represent a change in the criteria by which observers...

Figure 3. (overleaf) RT patterns in Experiment 2. Top panel: median RT for each shadow condition at each session, averaged across participants, orientation, target presence, and set size. Lower panels (one per participant): difference between median upright and median inverted RTs for each stimulus type at each session, averaged across target presence and set size. Leftmost sessions (1–12) are the initial learning trials. Rightmost sessions (GP and UL only) are the transfer trials immediately following stimulus change, represented by the vertical line. Light gray lines/bars represent 11%, mid-gray 20%, and black 45% contrast shadows; dotted line/striped bars, white non-shadows. Asterisks denote the significance level for the orientation × stimulus type interaction at each session for 20% contrast shadows vs. white non-shadows, calculated in a 2 (orientation) × 2 (stimulus type) × 2 (target presence) × 3 (set size) ANOVA for RTs to individual trials (**p < 0.01; *p < 0.05; *p < 0.1).
distinguish shadows from non-shadows, possibly compatible with Lovell et al.’s (2009) shadow-filtering system. **Experiment 2** did not allow us to distinguish these interpretations, although if adjusted criteria were shape-specific they should have failed to transfer to a new type of shadow stimulus. Nevertheless, these results at least allow us to exclude the primary visual cortex as a key player in discarding shadow-related information.

**Experiment 3: Fixation**

**Experiment 1** indicated that the slowing specific to upright shadows was associated with an increased number of saccades. One possible cause of this may be that shadows are more difficult to see in the visual periphery, thus more saccades are needed, to give more foveal input when attempting to distinguish shadow-like items. Indeed, 

Braje (2003) found that illumination direction affected recognition of centrally, but not peripherally, presented faces, suggesting that illumination-variant input like shadows may be more easily accessible in foveal areas as compared to the periphery. In neuronal terms, such a peripherally biased shadow suppression mechanism would show processing properties usually dedicated to higher level visual areas (at the level of object processing rather than feature processing), namely a non-uniform distribution of visual properties across the visual fields. Low-level visual areas, in contrast, are usually assumed to reveal a quite uniform distribution of their properties (Grill-Spector & Malach, 2004).

If shadow processing does differ in central and peripheral viewing, one would expect that preventing participants from moving their eyes should impair their access of shadow information, owing to more emphasis on peripheral visual input. To test this assumption, in **Experiment 3** we compared a free-viewing condition with one where fixation was maintained and the judgment was therefore conducted in more peripheral vision. We predicted that an upright–inverted RT difference would be greater with fixation, if shadow information was indeed less accessible in peripheral areas. In addition, we manipulated the distance of the target from central fixation: if shadow-related slowing was greater in peripheral vision, we would expect a greater (upright–inverted) RT difference for outer than inner targets, particularly under fixation conditions. This would be over-and-above the eccentricity effect usually described for visual search (Carrasco, Evert, Chang & Katz, 1995; Carrasco & Frieder, 1997) whereby, for stimuli of a consistent size, items near fixation and so of greater cortical magnification are searched faster than more peripheral items.

We also investigated whether shadow-specific slowing interacted with the side (left or right) on which the shadow is presented relative to the shadow caster. In our first experiments (and seemingly also those of Rensink & Cavanagh, 2004; Elder et al., 2004), the key shadow-like part of the stimulus always fell toward its left when inverted but to the right when upright. This could contribute to a disadvantage for upright shadows in two ways. Firstly, leftward perceptual biases have been reported (see review by Jewell & McCourt, 2000) with people attending more to the leftward than rightward parts of objects (Nicholls & Roberts, 2002; Orr & Nicholls, 2005; Post, Caulfield, & Welch, 2001). Secondly, it has been suggested that above left is the most likely assumed light source position (see McManus, Buckman, & Woolley, 2004; Sun & Perona, 1998), and hence shadows to the right are likely to be the most strongly discarded. In **Experiment 3**, we therefore counterbalanced shadow side for all stimulus conditions, allowing a comparison of the effects of shadows on the left with shadows on the right, for upright versus inverted stimuli. This also removed the need for a control based on shadow color. We hypothesized that any attentional bias would emerge in a main effect of shadow side, while any influence of light source position would result in an interaction of shadow side with stimulus orientation.

**Methodology**

**Participants**

Twelve psychology students (two males) aged 18–23 (mean 19.4) years participated in exchange for course credits. Two further subjects were tested but unable to perform at above chance level when fixating. As before, all gave written consent and the study was approved by the University of Bristol’s Faculty of Science Ethics Committee.

**Stimuli and tasks**

The basic task was identical to that used in **Experiment 1**, but all stimuli involved the white shadow casters with 20% contrast dark shadows, and their spatial presentation was more tightly controlled. A new set of 16 possible item locations was used throughout, the centers of which formed two concentric rings of diameters 6.3 cm and 11.7 cm around a central fixation cross, each ring comprising eight equally spaced locations. Individual stimulus dimensions and luminances approximated to those in **Experiment 1**. In each display, half of the shadow-casting elements were centered on locations in the inner ring and half on locations in the outer ring. The target (if present) was equally likely to occur in the inner and outer rings for each set size. Within these constraints, element location was randomly chosen on each trial. Both upright and inverted conditions included displays with shadows all to the left of the shadow caster and displays with shadows all to the right, in equal number. Only two set sizes (six and twelve items) were used.
Procedure

Each participant performed the task both with free eye movements and while fixating the central cross throughout, with order of the two conditions counterbalanced across individuals. Throughout both conditions, the eyes were monitored with the Eyelink II eye tracker. In the fixation condition, any trial on which the dominant eye’s gaze moved more than 1.5° from the center of the fixation cross during display presentation was aborted, to be replaced later in the sequence. In the free-viewing condition, the fixation cross was removed during display presentation and eye movement patterns were recorded for later analysis.

The test session began by showing the participant a sample stimulus display in order to explain the task. The eye tracker was then fitted. A practice block of 16 trials of the first condition was conducted, followed by two blocks of the same condition, each of 100 trials. The second condition was then introduced and a similar practice and two test blocks were completed. Each block was preceded by a calibration and validation sequence, and each trial by fixation of the central cross. The first four trials of each test block were omitted from analysis, and the remaining 96 comprised three of each combination of target presence, item orientation, set size, shadow side, and target eccentricity, presented in a randomized order. In total, each participant completed 24 target-present and 24 target-absent trials of each orientation at each set size for each viewing condition (free viewing or fixation), giving 384 trials.

Analysis

Percentage accuracy was calculated for each combination of viewing condition, orientation, target presence, set size, and shadow side, together with median correct RTs. Target present data were further sub-divided by inner and outer targets. The RTs were entered into an overall ANOVA, looking specifically for significant differences by orientation (slower responses to upright stimuli) and interactions of orientation with viewing condition (free viewing or fixation), shadow side, and target eccentricity. Eye movement variables (mean first and later fixation durations and number of saccades) for the free-viewing blocks were likewise examined to test for replication of the patterns found in Experiment 1.

Results and discussion

Accuracy was high for both viewing conditions and did not differ significantly (5.1% errors with fixation and 2.8% for free viewing). Although this difference, when taken with the RT data, hints at a speed-accuracy trade-off, any such effects were small.

An initial ANOVA was conducted on the median RT data using viewing condition, orientation, target presence, set size, and shadow side as repeated measures and testing order (free viewing or fixation first) as an independent measure. Since order showed no significant effect, the analyses were rerun excluding this variable. The expected main effect of item orientation was found ($F(1, 11) = 7.11$, $p = 0.022$; upright, 864 ms vs. inverted, 785 ms), together with a main effect of viewing condition ($F(1, 11) = 5.66$, $p = 0.037$): responses were slower with eye movements than without (926 ms vs. 724 ms). However, the interaction of viewing condition and orientation was non-significant ($p = 0.217$), suggesting that upright shadows are no more disadvantaged under fixation conditions. Separate ANOVAs for each viewing condition confirmed that the orientation effect was similarly marked in the fixation data ($F(1, 11) = 13.0$, $p = 0.004$; upright, 793 ms vs. inverted, 731 ms) and the free-viewing data ($F(1, 11) = 6.78$, $p = 0.025$; upright, 1023 ms vs. inverted, 912 ms). Thus, although eye movements slow the decision down, this is unrelated to shadow-like status. If shadow visibility (versus non-shadows) does not differ in fixation and free-viewing conditions, this also suggests that it does not differ in central and peripheral vision.

Whether shadows were on the left or right had no effect upon RTs ($p = 0.127$), nor did this interact with orientation ($p = 0.930$). We can therefore reject any strong influence of a leftward attentional bias and also of a predisposition to expect a top-left lighting context, upon the consistently reported slowing with upright shadows.

Repeating the analysis with just the target-present data and including target eccentricity as a variable, we found these same basic patterns plus a classic eccentricity main effect ($F(1, 11) = 18.3$, $p = 0.001$). Inner targets were identified faster than outer (698 ms vs. 793 ms). However, this did not interact with stimulus orientation ($p = 0.146$) nor further with viewing condition ($p = 0.238$). The consistency of the orientation effect for inner and outer targets fits well with the similarity of the effect under fixation and free-viewing conditions: shadow-related slowing seems to occur to a similar extent in peripheral vision and foveal areas. The greater number of saccades in shadow than non-shadow search cannot, therefore, be explained by reduced shadow visibility in the visual periphery nor are we able to narrow down the location of shadow-related slowing mechanisms to higher or lower level visual areas. While this result contrasts with the findings of Braje (2003), there are many differences between her study and ours, most notably perhaps her use of face stimuli, known to involve a contrasting central/peripheral processing bias from wider scenes (Grill-Spector & Malach, 2004).

Analysis of the eye movements showed a direct replication of Experiment 1: more saccades were made when searching upright than inverted items ($F(1, 11) = 7.91$, $p = 0.017$; mean 5.52 vs. 5.09, respectively). As before, first fixation durations did not differ by item orientation ($p = 0.567$), and later fixation durations, while approaching a significant orientation effect ($p = 0.064$),
followed the opposite pattern to that required to explain the RT differences (upright, 262 ms; inverted, 271 ms). Thus, while the RT data provide no explanation for why saccade numbers should increase when searching shadows, the eye movement data confirmed this to be the case.

**Experiment 4: Bilaterality**

In Experiment 3, we confirmed that, if free viewing is allowed, shadow search results in more saccades than non-shadow search. We also established that the shadow-specific slowing effect is unchanged if eye movements are eliminated and items viewed peripherally, meaning that we cannot exclude the involvement of low-level visual areas in driving shadow-related slowing. To explore this issue further, we then asked whether shadow processing requires the collaboration of both left and right visual fields and thus both hemispheres, or whether input to one visual field/hemisphere would be sufficient. It is generally accepted that low-level visual areas in the two hemispheres receive input from the collateral visual field only (Felleman & van Essen, 1991; Tootell, Mendola, Hadjikhani, Liu, & Dale, 1998). Therefore, if low-level visual areas were sufficient for shadow-related slowing, unilateral presentations of shadow displays should show the same difference between upright and inverted displays as the bilateral presentations used previously. In contrast, if shadow-related slowing affected bilateral presentations more than unilateral ones or even exclusively affected bilateral presentations, this would provide strong evidence that the mechanism depends upon higher level cortical areas, involving input from both hemispheres together (Liu, Zhang, Chen, & He, 2009). This idea was explored in Experiment 4.

As Experiment 3 had revealed a strong eccentricity effect, we adjusted stimulus size at the different eccentricities to compensate for cortical magnification (Carrasco & Frieder, 1997) in Experiment 4.

**Methodology**

**Participants**

Twelve students (two males) aged 18–38 (mean 20.8) years participated in exchange for course credits. Three further subjects did not complete the test session due to difficulties in fixating. As before, all gave written consent and the study was approved by the Faculty of Science Ethics Committee.

**Stimuli and tasks**

The fixation version of the task used in Experiment 3 was adapted in order that stimulus presentation could be restricted to one quadrant of the visual field (upper, lower, left, or right). The inner display ring (diameter = 6.3 cm) was divided into ten equally spaced possible stimulus locations and four of these were used, all adjacent and symmetrical about the horizontal axis for left/right displays or the vertical axis for upper/lower displays. Similarly, the outer ring (diameter = 11.7 cm) was divided into fourteen possible locations of which six were used. Thus, ten stimulus locations were available for each display position, and set sizes of six and ten items were presented. With six items, half appeared in inner locations and half in outer locations. Since ten items required all locations to be used, stimuli were presented with a randomized jitter of ±6 pixels in each direction to reduce the perceptual regularity of the display. The target, when present, appeared equally often in inner and outer locations. In addition, stimulus size was adjusted according to cortical magnification, so that the shadow caster for inner stimuli measured approximately 12 mm tall and for outer stimuli, 18.5 mm (see Carrasco & Frieder, 1997).

**Procedure**

Participants fixated the central cross throughout, and each completed one block of the task with the display in each position relative to fixation. Order of the four display positions was counterbalanced across individuals and always alternated between unilateral (left/right) and bilateral (upper/lower) positions. As in Experiment 3, the Eyelink II was used to monitor for fixation failures and reject and replace any such trials.

Each participant was first introduced to the task using the Experiment 3 displays, and allowed to practice the basic discrimination. The experimenter then explained the fixation requirements, the variable display position, and cortical magnification of the stimuli. The eye tracker was fitted and calibrated, and a practice block of the first condition was conducted (16 trials). One block of each of the four conditions then followed, each comprising 100 trials exactly as in Experiment 3.

**Analysis**

Accuracy and median correct RT data were compiled as in Experiment 3 but with viewing condition being replaced by display position, comparing unilateral (data for left and right visual field searches combined) with bilateral presentation (data for separate upper and lower visual field searches combined).

**Results and discussion**

An ANOVA using median RT data and comparing display position, orientation, target presence, set size, and shadow side demonstrated that the usual orientation effect was again present ($F(1, 11) = 11.2, p = 0.007$; upright, 1034 ms vs. inverted, 899 ms). However, display position...
had no overall effect \((p = 0.224)\), nor did it interact with orientation \((p = 0.497)\). As in Experiment 3, shadow side relative to the shadow caster (left or right) had no effect \((p = 0.584)\).

The same analysis with just the target-present data showed any effect of target eccentricity to have been eliminated by cortical scaling \((p = 0.843)\), indicating that this manipulation equated visibility of the stimuli. The interaction of eccentricity with orientation was non-significant \((p = 0.841)\) and did not further interact with viewing condition \((p = 0.894)\), confirming shadow-related slowing to be equivalent at both eccentricities tested.

The target-present analysis revealed a significant main effect of display position upon target identification speed \((F(1, 11) = 5.19, p = 0.044)\), responses being slower to bilateral (903 ms) than unilateral (818 ms) displays. Most importantly, there was a tendency for the orientation effect \((F(1, 11) = 17.1, p = 0.002)\); upright, 955 ms vs. inverted, 766 ms) to differ in bilateral and unilateral displays \((F(1, 11) = 4.39, p = 0.060)\). As the key question was whether or not unilateral presentations weakened the interpretation of a scene as containing shadows, we ran post-hoc analyses on this effect. Post-hoc Fisher’s LSD tests confirmed a greater orientation effect for targets in bilateral displays \((p = 0.005)\) than unilateral \((p = 0.034)\), see Figure 4a. Thus, bilateral display presentation seemed indeed to cause more problems with target identification than unilateral displays, particularly for upright shadow stimuli.

Accuracy was poorer and more variable in this than previous experiments, although all participants achieved over 80% correct overall. The patterns followed those for RTs, with errors biased toward target omissions when upright displays were presented bilaterally (see Figure 4b). The reported RT patterns could therefore be expected to be more marked if performance accuracy could be maintained.

The observation that shadow-related slowing occurs particularly with input from both visual fields suggests that the involvement of higher level cortical areas, receiving input from both hemispheres, is important in shadow-related slowing, and lower level mechanisms alone cannot be responsible (Liu et al., 2009). Somewhat counterintuitively, by presenting the display to one visual field/hemisphere only, the shadow task became easier, presumably because the shadow regions were more visible when this higher level input was eliminated.

Apart from providing further insight into the neural mechanisms underlying shadow processing, the above experiment can be understood as a manipulation of the contextual lighting framework within a visual scene. Scenes classified as containing shadows under bilateral presentation conditions might not have had sufficient contextual information to be classified as such under unilateral conditions. Such an interpretation is directly based on the anchoring hypothesis proposed by Gilchrist et al. for lightness and brightness perception (e.g., Gilchrist et al., 1999; for a review, see Gilchrist, 2006), an issue that we will consider further in the General discussion section.

### Experiment 5: Context

Experiments 1–4 established that shadow-related slowing occurs equally in central and peripheral vision, more with bilateral as compared to unilateral visual input, can be overcome by learning that is transferable, and, although not requiring eye movements, results in more of them if they are allowed. These findings indicate that both low- and higher level visual functions might contribute to shadow-related slowing. The question remains whether even higher level interpretive processes might further influence shadow-related slowing. In Experiment 5, we therefore examined the role of displaying the stimuli within an explicit lighting context, as a higher level interpretive input.

Specifically, we considered the implicit assumption that scenes are likely to be lit by a single illuminant from

![Figure 4](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933044/ on 11/21/2018)
above unless there is contradictory information (Adams, 2007; Ramachandran, 1988). This is integral to explaining the upright–inverted shadow differences reported (our Experiments 1 and 2; Elder et al., 2004; Lovell et al., 2009; Rensink & Cavanagh, 2004). One possible explanation for the reduced effect with unilateral viewing in Experiment 4 may be that shadows are only seen as such in the context of a consistent lighting direction, and such contextual information cannot be retrieved with unilateral viewing. In Experiment 5, we investigated one particular aspect of context by providing participants with an explicit and unchanging lighting scenario, to see if this modulated shadow-related slowing. It is known that people can learn to adjust their light-from-above assumptions according to experience (Adams et al., 2004). Here we tested whether a consistent lighting context different from the usual light-from-above scenario would reduce the extent to which our upright stimuli seem shadow-like and hard to perceive and conversely increase the shadow-like qualities of inverted stimuli. We hypothesized that the upright–inverted shadow difference may be greater for light-from-above than these other lighting scenarios, which we will call “light from below” even though this is not meant literally. In addition, we considered the congruency of the lighting direction with the shadows, and to this end included shadows falling in front of the shadow casters as well as shadows behind. If the shadow effect were context-specific and sensitive to detailed visual information such as congruency, lighting direction and congruency should interact, with the strongest slowing effect for upright shadows congruent with their lighting-from-above context.

Methodology
Participants
Ninety psychology students undertook the experiment as part of a first year laboratory class. Participation preceded any detailed information about the background or purpose of the experiment. The experiment was approved by the Ethics Committee of the Faculty of Science, University of Bristol, as before. Data from seven participants who had followed an incorrect testing procedure were excluded, as were those from a further seven whose overall error rates were more than 2 SDs above the mean. The final sample consisted of 76 participants (10 males) aged 18–47 (mean 19.6) years.

Stimuli and tasks
The task was based upon the free-viewing condition of Experiment 3, but with modifications to the overall context in which the search array was displayed (see Figure 1d). The two rings of stimulus locations, now of diameters 10.8 cm and 16.4 cm around the fixation cross, were centered in a clearly defined circular region of the display, of diameter 21.4 cm. Individual stimulus dimensions approximated to those in Experiment 1. The luminance of this display area was measured as 42.1 cd/m², with the shadows recording a Michelson contrast of 20% relative to this. The surrounding screen was black (luminance = 2.20 cd/m²), other than a region defined by tangents to the display area, which was lighter than the background (22.3 cd/m²) and intended to suggest illumination of the display area by a light beam off-screen. The central axis of the “light beam” was oriented at 30° to the vertical. Light source position was manipulated as a between-subjects variable, with four conditions included: lighting from above right (final n = 22), above left (n = 16), below right (n = 18), or below left (n = 20). Whichever lighting condition was involved, the contextual background was present throughout all test blocks.

In addition to shadows apparently behind the shadow caster (as in Experiments 1, 3, and 4), we included displays with stimuli for which the shadow was apparently in front, as for the final trials of Experiment 2. Shadow location (behind/in front) was the same for all items in a display but varied between displays. Each test condition involved trials with upright stimuli (shadows at the bottom) and inverted stimuli (shadows at the top) for both shadows in front and behind and shadows to the left and right. As in Experiments 3 and 4, set sizes of six and twelve items were used.

Procedure
Half of the group was tested simultaneously in a dimly lit computer room, followed immediately by the other half. Each individual completed five sequential blocks of testing, each block involving 64 trials of the task, the first block being discarded as a practice leaving 256 trials for analysis. Displays were presented on 17" LCD monitors viewed at 57 cm. When ready to proceed, the participant was instructed to press any key to initiate a stimulus display, and then the appropriate response key immediately when the decision was made. The fixation cross appeared after a pause of 1000 ms, accompanied after a further 1000 ms by the stimulus display for the next trial. On response, the display cleared, leaving just the contextual background. Each test block comprised one target-present and one target-absent trial for each combination of item orientation, set size, shadow side (left/right), shadow location (behind/in front), and target eccentricity, presented in a randomized order.

Analysis
Percentage errors and median correct RTs were calculated across equivalent trials for each individual. All variables (except, initially, target eccentricity) were entered into an overall mixed ANOVA, with lighting condition as an independent variable and orientation,
target presence, set size, shadow side, and shadow location as repeated measures.

Results and discussion

As in all previous experiments, upright shadows were searched more slowly than inverted (1.29 s vs. 1.16 s; \(F(1, 72) = 58.3, p < 0.001\)). Error rates averaged 5.3% across all conditions and followed similar patterns to RTs.

Lighting condition showed no overall main effect on RTs (\(p = 0.998\)), no interaction with orientation (\(p = 0.930\)), and no interaction with any other factor except target presence (\(F(3, 72) = 0.047\)) for which no post-hoc comparisons approached significance. Although any patterns related to lighting congruency should have emerged in a complex lighting condition \(\times\) shadow side \(\times\) shadow position interaction, we further checked for any congruency effects by categorizing shadow directions into “congruent” and “incongruent” with each of the four lighting conditions and collapsing across horizontal lighting direction (i.e., only “light from above” and “light from below” were distinguished), target presence, and set size. We first took “congruent” shadow directions as lying parallel to the central axis of the light source, irrespective of the shadow caster position. No effects involving congruency approached significance. We then narrowed the definition of “congruent” to include only those of the displays previously identified for which the shadow caster fell between the light source and the shadow, all other displays being classed as “incongruent”. Again congruency had no significant effects, and we conclude that explicitly knowing the light source direction, as manipulated in this experiment, did not influence the slowing effect associated with upright shadow stimuli. Possibly an effect would have emerged with more prolonged experience (note that many trials are needed to overcome the upright–inverted difference in Experiment 2), and/or more powerful conflicting information, such as provided by haptic feedback of shape or position (see Adams et al., 2004). As further evidence of the resistance of the upright–inverted effect to obvious lighting changes, the patterns were unaffected by whether shadows fell behind or in front of the shadow caster (\(p = 0.314\)). Nevertheless, even if an explicitly depicted lighting context as used here does not readily affect shadow perception, we cannot conclude that implicit contextual or scene-based information is not important in shadow processing.

Although in previous experiments shadow side was not a significant modulator of the results, here search was slightly faster overall with shadows to the left than to the right (1.20 s vs. 1.25 s; \(F(1, 72) = 21.1, p < 0.001\)), in keeping with the literature suggesting a leftward object-based attentional bias (Jewell & McCourt, 2000; Orr & Nicholls, 2005; Post et al., 2001). This interacted with orientation and target presence (\(F(1, 72) = 4.12, p = 0.046\)), the left–right difference being greater for upright targets (\(p < 0.001\), Fisher’s LSD) than inverted (\(p = 0.014\)) or target-absent displays (\(p > 0.4\); see Figure 5a). This may suggest that the light-from-above shadow-processing disadvantage is more pronounced with shadows on the right, in keeping with the idea that a light-from-above-left assumption dominates perception (McManus et al., 2004; Sun & Perona, 1998) and therefore right-side shadows are least well processed. The interaction of the shadow inhibition effect with the leftward attentional bias might also indicate that the two share common processing levels. However, given that this lateralization effect did not emerge in Experiments 3 and 4, any such effect is apparently small (note though that the larger sample in Experiment 5 gave more power).

Finally, the target-present RT data were analyzed separately to check for eccentricity effects. As expected (Carrasco et al., 1995; Carrasco & Frieder, 1997), inner targets were found faster than outer (972 ms vs. 1114 ms; \(F(1, 65) = 60.1, p < 0.001\)), but this also interacted with shadow orientation (\(F(1, 65) = 14.5, p < 0.001\)) even if \(z\)-scored for inner and outer targets separately (interaction, Figure 5).

Figure 5. Orientation interactions in Experiment 5 (target-present data only). (a) Orientation \(\times\) shadow side. (b) Orientation \(\times\) target eccentricity. Details as for Figure 4.
Cast shadows are difficult to perceive, presumably in order to minimize noise deriving from irrelevant visual information and thus facilitate object recognition (see Elder et al., 2004). Regarding the underlying shadow-processing mechanism, one suggestion is that systems in early sensory areas may identify and discount shadows, resulting in their being difficult to access, with attention presumably required for their retrieval (Rensink & Cavanagh, 2004). Evidence for such an assumption comes from the observation that visual search for shadow-like items is slowed relative to the rapid search for identical items that are inverted and therefore not perceived as shadow-like (Rensink & Cavanagh, 2004). However, since cast shadow information can help with object recognition or understanding the spatial layout of visual scenes (Allen, 1999; Castiello, 2001; Hubona et al., 2004; Tarr, Kersten, & Bülthoff, 1998), it cannot be entirely discarded.

An alternative proposal is that an early, coarsely scaled shadow filter might separate shadow information from non-shadow information (Lovell et al., 2009). Here, we tried to shed new light upon the neural mechanisms that might underlie shadow processing and access to shadow information.

In five visual search experiments, all based upon Rensink and Cavanagh’s (2004) original paradigm with cast shadows, we collected evidence that speaks against the existence of an early shadow-specific processing stage (Lovell et al., 2009) and that suggests that higher level rather than low-level sensory mechanisms are the key players in shadow inhibition. Firstly, slowing for search tasks requiring fine-scaled shadow (relative to non-shadow) discrimination was not reflected in prolonged first fixation durations in Experiment 1, as would have been expected if a preliminary shadow-filtering stage was involved (Lovell et al., 2009). Secondly, with training, participants were able to search upright shadow-like stimuli as fast as non-shadow stimuli (Experiment 2), a learning process that completely transferred to a different form of shadow stimulus. Since learning one low-level sensory discrimination does not help with performing another low-level sensory discrimination (Ahissar & Hochstein, 1993; Sireteanu & Rettenbach, 1995, 2000), these results are thought to exclude the involvement of such low-level sensory areas as V1. Thirdly, shadow-related slowing effects were weakened by unilateral as compared to bilateral display presentation, indicating that low-level visual areas on their own were not sufficient to induce shadow-related slowing. However, shadow-related slowing effects did not differ in central from peripheral visual areas (Experiments 3 and 4, but see Experiment 5) and hence were equally present under fixation and free-viewing conditions (Experiment 3). These findings make it impossible to exclude an involvement of low-level visual areas in shadow processing.

We also established that, at whatever levels the slowing occurred, it was manifested in an increased number of eye movements, typically supposed to indicate a change in search strategy and therefore decision-related high-level processing (Zelinsky & Sheinberg, 1997). Likewise, the ability to learn to overcome shadow-related slowing (Experiment 2) could reflect the learning of better strategies (in addition to improved attentional deployment). However, the extra saccades were not seemingly required for additional checking of the target, nor because the fixation of shadowy image regions was inhibited (as indicated by point-of-gaze analysis), nor because shadows were less visible in peripheral than foveal areas. Exploration characteristics did not otherwise differ from the eye movements made during non-shadow search and the error rate did not increase for shadow search as compared to non-shadow search. These observations seem to exclude the use of different search strategies for shadows, due either to greater uncertainty in localizing a shadow-like target or to different decision criteria. Thus, just as low-level sensory processes do not seem to be sufficient to drive shadow-specific slowing, voluntary decision-related processes are not.

Further than this, we can only speculate about the actual mechanism leading to shadow-specific slowing. If neither low-level sensory areas are sufficient nor high-level (decision-making related) strategies, shadow-related slowing might be primarily driven by cortical areas somewhere between these two extremes.

Even though an admittedly small effect, the observed interaction between shadow-related processing and object-based attentional biases in Experiment 5 hints toward object recognition as a crucial level at which differences between upright and inverted shadow-like items may emerge. Such an idea is not new—it was raised before by Elder et al. (2004). However, why should object recognition be at the core of shadow-related slowing? Here, the concept of combined heuristics for distinguishing “lighting” from “material” (Kingdom, 2008) may provide a framework within which to understand these effects. One might speculate that, in the upright shadow condition, such heuristics lead us to identify only the shadow casters (i.e., vertical bars) as “material” (i.e.,
objects), while the shadows (i.e., oblique bars) would be classified as lighting, not contributing to object recognition. Any automatic search based on differences between target and distractor objects would then fail, since no differences exist at this level of discrimination. In other words, the readily accessible high-level object information is insufficient for the task, and attention is required to additionally access the lower level orientation information and allow the shadow differences to be identified. In the terms of a reverse hierarchy (Hochstein & Ahissar, 2002), feedback connections are required to allow reentry to the low-level feature information. In the inverted condition, in contrast, the same heuristics would identify both vertical and oblique bars as “material” of different surface structure, together forming an angular object. Angular target objects differ from angular distractor objects, allowing search to be based on discrimination of high-level object information, which is readily accessible via feedforward processing mechanisms (Hochstein & Ahissar, 2002). This has much in common with Rensink and Cavanagh’s (2004) idea that information leading to shadow identification is processed via information circulating between various levels within the visual system. What might be interpreted as an early discarding of shadow information requiring later retrieval could similarly be understood as the consistent availability but reduced accessibility of shadow information to high-level object processing mechanisms if somehow classified as lighting. Thus, the suppression of shadows might occur as a normal component of object processing, rather than requiring a discrete shadow-specific processing system.

If this speculation is correct, then whether or not slowed shadow-specific processing occurs would depend upon how people use multiple cues in the natural environment to decide what is material (i.e., objects) and what is lighting-related (Kingdom, 2008). Perceptual learning might allow participants to readjust how such cues are interpreted, perhaps by adjusting their Bayesian priors for defining material versus lighting so that the shadow-like part of the upright search items becomes perceived as material and not lighting. Such an interpretation of our results would also allow us to account for the dependence of shadow-related slowing on specific stimulus presentation factors. The implicit rules for combining cues and heuristics appear complex and it is highly likely that the saliency of each varies with the context provided by other parts of the visual input (see Gilchrist, 2006; Gilchrist et al., 1999). Thus, the increased shadow visibility under fixation conditions with unilateral displays (Experiment 4) as compared to bilateral displays could be because insufficient higher level contextual information was available from just one visual field to allow the shadow-like part to be attributed to lighting, and thus suppressed. The slight exaggeration of shadow-related slowing for peripheral over central and right-side over left-side stimuli in Experiment 5 could be because these regions, being of lower processing priority (Carrasco et al., 1995; Jewell & McCourt, 2000), suffer more from the need for attention to access the vital information and/or are processed later than inner/left-side items, which allows more time for the contextual effects to emerge before response. Although our explicit manipulation of contextual light source direction did not affect shadow-related slowing, this could have been because our images were too stylized to provide convincing lighting information. Lighting and an implicitly perceived light source may nevertheless be key components of the contextual information if presented more naturally.

Conclusions

What makes cast shadows hard to see? It seems likely that neither low-level sensory shadow discounting mechanisms nor the high-level strategies applied to exploring a visual scene are sufficient to induce shadow-related slowing. Instead, we suggest that the difficulty in accessing shadow-specific information is tied in with object recognition and in particular the parsing of visual input into physical objects and illumination-related noise. The enormity of this task and the complexity of the implicit rules applied go some way to explaining why shadow-related research has, to date, been somewhat inconclusive. Whether low-level mechanisms are at all involved in identifying shadows as shadows needs further investigation.

Acknowledgments

We thank Ron Rensink and an anonymous reviewer for their helpful comments on an earlier draft of this manuscript. This research was supported by a grant awarded to AT from the BRACE Charity (Bristol Research into Alzheimer’s and Care of the Elderly), as a basis for further work examining shadow processing in aging and Alzheimer’s disease.

Commercial relationships: none.
Corresponding author: Gillian Porter.
Email: gillian.porter@bristol.ac.uk.
Address: 12a Priory Road, Bristol BS8 1TU, UK.

Footnotes

1 All search slope analyses were conducted as described by Rensink and Cavanagh (2004). For combined search rates, individual target-absent slopes were divided by the overall ratio of group mean target-absent/target-present slopes for that condition. The mean of target-present and
weighted target-absent slopes for each individual was then analyzed.

A circular region of interest (ROI) was identified for each display, 2° in radius and centered on the vertical mid-point of the target’s shadow caster on its “shadow” side. This just encompassed the entire target shadow caster and “shadow”. Landing points of saccades relative to this ROI were analyzed.

References


