The perception of three-dimensional cast-shadow structure is dependent on visual awareness

Sieu K. Khuu
The School of Optometry and Vision Science, The University of New South Wales, Sydney, NSW, Australia

Jack Gordon
The School of Optometry and Vision Science, The University of New South Wales, Sydney, NSW, Australia

Kaleah Balcomb
The School of Optometry and Vision Science, The University of New South Wales, Sydney, NSW, Australia

Juno Kim
The School of Optometry and Vision Science, The University of New South Wales, Sydney, NSW, Australia

In the present study we examined whether the perception of depth from cast shadows is dependent on visual awareness using continuous flash suppression (CFS). As a direct measure of how the visual system infers depth from cast shadows, we examined the cast-shadow motion illusion originally reported by Kersten, Knill, Mamassian, and Bulthoff (1996), in which a moving cast shadow induces illusory motion in depth in a physically stationary object. In Experiment 1, we used a disparity defined probe to determine the stereo motion speed required to match the cast-shadow motion illusion for different cast shadow speeds (0°/s–1.6°/s) and different lighting directions. We found that configurations implying light from above produce more compelling illusory effects. We also found that increasing shadow speed monotonically increased the stereo motion speed required to match the illusory motion, which suggests that quantitative depth can be derived from cast shadows when they are in motion. In Experiment 2, we used CFS to suppress the cast shadow from visual awareness. Visual suppression of the cast shadow from awareness greatly diminished the perception of illusory motion in depth. In Experiment 3, we confirmed that while CFS suppresses the cast-shadow motion from awareness, it continues to be processed by the visual system sufficient to generate a significant motion after effect. The results of the present study suggest that cast shadows can greatly contribute to the perception of scene depth structure, through a process that is dependent on the conscious awareness of the cast shadow.

Introduction

The visual system provides valuable perceptual information about the depth order of objects and surfaces in the environment relative to the observer, which is essential for object localization and visually guided behavior. This task is not trivial; the depth structure of the scene is not mapped directly into 2-D retinal images, but must instead be inferred from a host of primary (e.g., binocular disparity) and secondary (e.g., size, perspective, and motion parallax) image-based cues that correlate with depth order. Although the visual system relies on binocular disparity as a metric cue to depth, it can also rely on a number of monocular cues, such as shadows, that provide an ordinal representation of object positions in depth. Shadows are regions of illumination change produced when an object obstructs light and are associated either with the object surface (attached shadow) or the background surface (cast shadow) and are ubiquitous in natural scenes (see Yonas, 1979).
Cast shadows are particularly informative about depth order because the position and shape of shadows cast on a projection plane cues the relative 3-D spatial position of objects in the scene. Kersten, Mamassian, and Knill (1997) give a powerful demonstration of this effect, showing that a moving cast shadow can be used to infer the direction of a ball moving through a virtual environment. In an analogous effect, a cast shadow undergoing rigid lateral 2-D motion can induce illusory motion in depth in a physically stationary object (Kersten, Knill, Mamassian, & Bülthoff, 1996). This cast-shadow motion illusion reveals that the visual system assumes that backgrounds (reflecting the statistical properties of natural scenes) are stationary and that the object is moving in depth, rather than that the light source is moving. This striking illusion demonstrates that the visual system, aided by prior assumptions of regularity in the visual scene, is able to use cast shadows as cues to the 3-D spatial layout of objects and the visual scene (see Allen, 1999).

Previous studies have sought to clarify how cast shadows are processed by the visual system. Some research has suggested that the processing of cast shadows is mediated by a coarse spatial-scale analysis of the visual scene in which objects and shadows are matched based on the global properties of the stimulus (such as the center of mass; see Mamassian, 2004). Matches at a local level are largely ignored. This coarse-scale analysis provides an efficient solution to the shadow correspondence problem and a means of quickly establishing the direction of illumination to determine the 3-D spatial layout of the visual scene. The detection of cast shadows on a coarse spatial scale accounts for both the relative insensitivity for illumination change detection in the perception of cast shadows that has been noted by a number of previous studies (e.g., Enns & Rensink, 1990; Jacobson & Werner, 2004; see also Khuu, Khambiye, & Phu, 2012; Khuu, Moreland, & Phu, 2011), and the observation of “impossible shadows,” common in art where 3-D structure is still apparent despite local inconsistencies in the illumination direction (see Farid & Bravo, 2010; Mamassian, 2004).

The fact that 3-D structure is readily evident from cast shadows (e.g., the cast-shadow motion illusion) has naturally led to the assumption that their processing might occur implicitly without the need for attention and visual awareness. That is, the visual system does not have to be conscious of cast shadows in the environment for them to influence perception. Whether cast shadows are implicitly or explicitly processed by the visual system is presently unresolved and only a handful of studies have sought to address this issue (see Dee & Santos, 2011, for a review). Currently, the consensus view is that the processing of cast shadows is implicit, automatic, and occurs in comparatively earlier stages of image analysis (see Rensink & Cavanagh, 2004). Research supporting this view comes from visual search studies in which cast shadows are used to cue a target element amongst a field of distracters that have cast shadows with a different orientation from the target (e.g., Cunningham, Beck, & Mingolla, 1996; Elder, Trithart, Pintilie, & MacLean, 2004; Lovell, Gilchrist, Tolhurst, & Trosianko, 2009; Rensink & Cavanagh, 1993, 2004). These studies have shown that the visual system is able to rapidly (as measured by reaction times) and implicitly identify the target cast shadow amongst a field of distracters within a few 100 ms from stimulus onset, as measured by reaction times. This suggests that cast shadows might be analyzed at early stages of visual processing without the need for visual awareness (see Rensink & Cavanagh, 2004). This assertion is in agreement with the findings of Harris, Schwarzkopf, Song, Bahrami, and Rees (2011) who demonstrated that other surface properties produced by the interaction between an object and a light source, such as apparent brightness and contrast, are determined without visual awareness. Both Rensink and Cavanagh (2004) and Casati (2006) have further proposed after this initial detection stage, cast shadows are “ignored” or disregarded to prevent them from interfering with the detection of objects in the visual field. However, recent studies (e.g., Khuu et al., 2012; Lovell et al., 2009; Mamassian, 2004) have suggested that human insensitivity to cast shadows might reflect a coarse scale process is solving the shadow correspondence problem, which largely ignores fine spatial detail.

Additional evidence for the implicit processing of cast shadows comes from Castiello, Lusher, Burton, and Disler (2003; see also Castiello, 2001), who compared the effect of cast shadows on the recognition of objects in normal observers and those with visual neglect (a disorder of visual attention) arising from hemispherical damage. Patients whose visual neglect did not arise from damage to the right temporal lobe were able to use cast shadows without awareness to facilitate the recognition of objects to the same extent as normal observers. These findings are consistent with the view that the identification of cast shadow occurs implicitly and without the need for observers to be explicitly aware of the shadow.

Importantly, the abovementioned studies have drawn their conclusions by examining how cast shadows contribute to the identification and recognition of target objects. For example, because cast shadows are frequently darker backdrops to objects, they might serve to enhance object contrast, allowing the object to stand out from its background, which facilitates recognition in these studies. However, these judgments represent only one aspect of how cast shadows contribute to perception. As noted, cast shadows are greatly informative about 3-D object
structure. To our knowledge, no study has examined whether the extraction of relevant 3-D information from cast shadows is dependent on visual awareness and how this process occurs.

Although the initial interpretation of cast shadows provides the basis for visual search and the recognition of objects, the inference of depth from cast shadows must depend on an additional process in which the spatial positions of the object and cast shadow are first estimated and then compared. The previously described cast-shadow motion illusion may affect this process, and the perception of illusory object motion arises through a temporal change in the apparent depth signaled by the moving cast shadow. Such a process most likely requires additional processing beyond simply identifying regions of visual space as cast shadows. Indeed, it has been argued that the inference of 3-D structure from depth cues is a complex process that involves the integration and perceptual grouping of local information in higher cortical areas (see Bulthoff & Mallot, 1988; Johnston, Cummings, & Landy, 1994; Lee, Khuu, Li, & Hayes, 2008; Swain, 1997), perhaps involving top-down processing (see Bulthoff, Bulthoff, & Sinha, 1998; Lamoureul, Cornilleau-Pérès, & Droulez, 1997). Recent neural imaging studies have identified specific areas high in visual processing (in particular areas middle temporal, MT, the bilateral caudal inferior temporal gyrus, and lateral occipital sulcus) that are responsible for the representation of 3-D structure through the integration of different depth cues such as motion, texture, stereo, and shading cues (e.g., Ban, Preston, Meeson, & Welchman, 2012; Georgieva, Todd, Peeters, & Orban, 2008; Liu, Vogels, & Orban, 2004). Though no imaging study has directly examined the neural processing of cast shadows, it could be argued that information from cast shadows might also contribute to processing in depth-sensitive cortical areas.

Although the inference of 3-D object structure from depth cues most likely reflects additional processing (beyond an initial identification process that occurs without awareness), it remains largely unclear whether this process (in relation to the processing of cast shadows) is dependent on visual awareness. Previous studies highlight that, while visual awareness is generally not a requirement for the processing of basic stimulus features such as local orientation and brightness, the perception of more complex stimuli, particularly those that require spatial integration and or grouping of local features, does depend on visual awareness, though the extent of this dissociation remains a topic of much empirical research (e.g., Lin & He, 2009; Tononi & Edelman, 1998, c.f., Mudrik, Breska, Lamy, & Deouell, 2011). As cast shadows are used by the visual system to infer the 3-D structure of objects, it is possible that this process might depend on visual awareness. The extent to which this is the case is the focus of the present study.

In this study, we investigated whether cast shadows enhance perception of 3-D structure through implicit (not requiring awareness) or explicit processing (requiring awareness). To this effect, we quantified the cast-shadow motion illusion (originally reported by Kersten et al., 1996), which provides a powerful demonstration of how cast shadows are used to infer the 3-D relationship of an object in the environment. We also examined whether binocular suppression of the moving cast shadow modulates the extent of perceived illusory motion in depth. In the present study we suppressed the moving cast shadow using interocular suppression, in particular a variant of continuous flash suppression (CFS) devised by Tsuchiya and Koch (2005), in which powerful binocular suppression is achieved by presenting the to-be-suppressed stimulus to one eye and a dynamic mask (a Mondrian field at 10 Hz) to the other eye. When viewed binocularly, the target stimulus is suppressed from view by the dynamic mask despite the fact that it is monocularly “visible” to one eye. Unlike traditional binocular rivalry, CFS achieves sustained and stable suppression (of the entire stimulus or parts of the stimulus) over much longer periods of time. CFS is therefore a very useful paradigm for investigating the limits of conscious vision (see Koch & Tsuchiya, 2007; Lin & He, 2009).

We predicted that if the extraction of 3-D structure from cast shadows is processed without awareness, suppressing them from view using CFS should not affect the strength of illusory motion. This is because the visual system has already extracted meaningful depth information from the stimulus early in processing, so awareness of the cast shadow is not required to perceive the illusion. Accordingly, the percept of illusory motion will be the same regardless of whether the observer’s awareness is modulated by CFS of the cast-shadow motion. Alternatively, if the inference of depth from the perceptual grouping of cast shadows and objects is an explicit process that requires conscious awareness, suppressing awareness of the cast shadow will attenuate, or abolish, illusory motion. In the present study, we asked which of these two outcomes accounts for perception, thereby assessing the role of visual awareness in the perception of depth from cast shadows. Three experiments were conducted. In Experiment 1, we directly quantified the cast-shadow motion illusion by matching its apparent speed to a stereo motion probe for different shadow speeds and illumination directions. In Experiment 2 we repeated Experiment 1 for conditions in which the cast shadow was made either visible or invisible using CFS. In Experiment 3, we performed an adaptation study to
verify that the effects were not due to CFS eliminating the cast shadow from visual processing altogether.

Experiment 1: Using a stereo motion probe to quantify the perceived motion in depth induced by a moving cast shadow

Determining whether visual awareness is a requirement for the perception of the cast-shadow motion illusion first requires a direct means of measuring the perceived depth of the stimulus. In the original study, Kersten et al. (1996) asked observers to report whether the stimulus appeared to move in depth for conditions in which either the shadow orientation or blur were varied. While this method is important in initially establishing the phenomenology of the cast-shadow motion illusion and the stimulus conditions under which its perception is optimal, it does not measure the extent of perceived motion in depth traversed by the stimulus. In Experiment 1, we addressed this issue by quantifying the speed of the cast-shadow motion illusion by directly comparing it with a reference stimulus whose motion in depth was defined by stereo motion produced by changing binocular disparity. Here, the goal was to quantify and determine the congruence in the perception of motion in depth facilitated by two different depth cues. We were particularly interested in the stereo motion speed required to match the speed of the cast-shadow motion illusion.

A number of different conditions were created to examine the effect of changing the speed of the cast shadow on the perception of illusory motion in depth. Faster cast shadow speeds might imply faster object motion. Consequently, increasing shadow speed might systematically change the extent of perceived motion in depth. In turn, this would suggest that cast shadows can provide quantitative information about the 3-D scene (as suggested by Allen, 1999), rather than simply an ordinal cue to depth. In addition to changing shadow speed, the position of the cast shadow was also changed to imply different lighting directions. A well demonstrated property of cast-shadow detection is a bias for a light source that is fixed and above (Kersten et al., 1996; Khuu et al., 2012; Ramachandran, 1988; Rensink & Cavanagh, 2004). Accordingly, stronger illusory effects might be expected for this “ecologically valid” configuration compared with lighting directions implying a source that is below or to the side.

Methods

Observers

Six experienced observers with normal or corrected-to-normal visual acuity participated in the experiment. One was the author (SKK), while the others were naive to the objectives of the experiment.

Stimulus

The stimulus was a movie sequence (duration of 1 s) presented monocularly to the left eye showing a uniform white square (64 cd/m², 2° × 2°) superimposed on a Mondrian surface. Across each frame transition, the square remains physically stationary, and the cast shadow undergoes lateral motion (to the lower right) on a Mondrian surface. Note also that the cast shadow becomes progressively blurred at greater separations to simulate a changing penumbra.

Figure 1. Static frames of the cast-shadow stimulus used in the present study. Across each frame transition, the square remains physically stationary, and the cast shadow undergoes lateral motion (to the lower right) on a Mondrian surface. Note also that the cast shadow becomes progressively blurred at greater separations to simulate a changing penumbra.
onto a square grayscale Mondrian (4° x 4°, luminance range: 16–48 cd/m², see Figure 1) field that changed randomly at 10 Hz (see Figure 1). The background was made dynamic so that its temporal profile was similar to CFS conditions used in Experiment 2 and to avoid local textures from being used as a reference to judge the physical position of the object. The flickering background did not impair the perception of illusory motion, which suggests that the visual system is able to segregate and distinguish the motion of the cast shadow from the moving background texture. The right eye always viewed a uniform gray field set to a background luminance of 20 cd/m². Thick black lines (6 cd/m², width: 0.25°, length: 7.5°) that bordered the stimulus area were presented to both eyes, and this area was additionally placed in the center of a larger static Mondrian field (8° x 8°) that was identical in both eyes. The black lines and static Mondrian field aided stereo fusion and minimized convergence. The white square was accompanied by its apparent cast shadow—a darker square (12 cd/m²) of similar size that was occluded, appearing beneath the stimulus. This resulted in a vertically aligned white square with its cast shadow projected onto a vertical background surface (Figure 1). To generate shadow motion, the cast shadow was displaced laterally by a fixed jump size across movie frames in a particular direction. Changing the jump size changed the apparent speed of the shadow, while the direction of motion provided an indication of the light source direction (see below). Note that as the shadow is displaced away from the stimulus, its edges were increasingly blurred (using a Gaussian filter in which the standard deviation of the filter was proportional to the shadow separation) to simulate an increasing penumbra. This change is a natural property of cast shadows to which the visual system is highly sensitive (Kersten et al., 1996); when incorporated into the cast shadow, it produces more compelling illusory motion in depth.

These stimuli were generated using custom software written in MATLAB (version 2013b) on a Macintosh iMac 2.8GHz computer. Stimuli were viewed on a linearized 23-in. “True3D” 3-D monitor viewed through polarized lenses at a viewing distance of 57 cm. The stimulus was presented 4° to the right from the center of display, which was marked by a white square that appeared throughout the experimental trials. Observers were instructed to fixate on this mark.

**Procedure**

The shadow stimulus was compared to an identical square stimulus whose motion was defined only by stereo motion. To generate this reference stimulus, pairs of the abovementioned stimuli were presented without a cast shadow and separately to the eyes. Stereo motion was produced by horizontally displacing the corresponding white squares in opposite directions. These stimuli were presented for 1 s in two temporal intervals, separated by an interstimulus interval of 0.5 s in which the screen was blank and set to gray (32 cd/m²). The shadow stimulus was presented in the first interval while the stereo motion stimulus was presented in the second interval. The task of the observer was to, while maintaining fixation, attend to the squares presented in the two intervals and judge the interval containing the fastest square moving in depth. A staircase (converging on the 79% correct performance level) was used to systematically change the stereo motion stimulus to perceptually match the cast-shadow stimulus. The initial stereo motion speed was randomized between 0.55/° s and 0.35/° s and the step size was 0.02/° s. On each subsequent reversal the step size was halved, and after the fourth reversal the step size remained at a value of 0.01/° s. Subpixel step sizes were achieved by using anti-aliasing procedures, which modulated the luminance of the edges of the cast shadow in accordance with its subpixel placement. The staircase lasted for eight reversals; the average of the last four reversals provided an indication of the stereo motion speed required to match the shadow stimulus. This procedure was repeated for six different shadow speeds of 0/° s, 0.1/° s, 0.2/° s, 0.4/° s, 0.8/° s, and 1.6/° s and for three illumination directions implying light from above at 45° to the left, light from below at 45° to the right, and light from the left side.

Observers performed 10 blocks of 18 staircase trials. Within each block the six different shadow speeds were repeated for the three different illumination conditions. The trial order was randomized within and between blocks. The perceived speed of the shadow stimulus (for each shadow speed and lighting direction) was determined by averaging the threshold value across the 10 blocks of trials.

**Results and discussion**

The results of Experiment 1 are plotted in Figure 2, which shows the average judged stereo motion speed (binocular disparity in degrees per second) required to match the cast-shadow stimulus, plotted as a function of shadow speed (degrees per second in log scale); error bars signify 95% confidence intervals. Different symbols represent data for the three different lighting conditions. A two-way repeated measures analysis of variance (ANOVA) was conducted to examine the effect of lighting direction (Factor 1, three levels) and shadow speed (Factor 2, six levels) on the stereo motion speed required to match the cast-shadow stimulus. This analysis confirmed main effects for both shadow speed, $F(5, 30) = 164.95, p < 0.0001$, and lighting direction,
Some further findings are worthy of note. First, we found that when the shadow was stationary (as expected) the matched stereo motion speed for this condition which was $0^\circ/s$ (open symbols), which indicated that observers did not report any illusory motion in depth. This is noteworthy because it shows that observers were able to accurately perform the matching task, and it provides a useful baseline for comparison with the moving cast-shadow conditions.

Second, the matched stereo motion speed increased monotonically with the speed of the cast shadow, though the extent of perceptual change was dependent on the lighting direction. This result conclusively demonstrates that the strength and extent of illusory motion in depth (as indicated by the matched stereo motion speed) increased with shadow speed. For the fastest shadow speed used in the present study (i.e., $1.6^\circ/s$) the stereo motion speed required to match the cast-shadow stimulus was approximately $0.12^\circ/s$. To provide an indication of the dependency of the cast-shadow motion illusion on the speed of the cast shadow, least squares linear regression analyses were conducted separately for the three lighting conditions, and the line of best fit (average $R^2$: 0.92) is shown in Figure 2A (gray solid lines). The obvious dependency of the motion illusion on shadow speed suggests that when in motion they powerfully contribute to the apparent depth of objects (see Dee & Santos, 2011; Kersten et al., 1996, Kersten et al., 1997). This observation is consistent with the findings of Allen (1999), who showed that cast-shadow properties such as angle of illumination, interposition, and position can all be used by the visual system to estimate the distance of objects from the a background surface.

Third, matched stereo motion speed was dependent on lighting direction. The stimulus configuration in which light is implied to be above produced stronger illusory effects, as indicated by the slope of the best fit line (slope: $0.074 \pm 0.0075$, 95% CI), than when light was from left (slope: $0.034 \pm 0.0052$, 95% CI) and from below (slope: $0.028 \pm 0.0047$, 95% CI); for these two lighting directions the cast-shadow motion illusions were similar and not significantly different from each other. Indeed, Tukey’s Honesty Significant Difference (HSD) post-hoc tests (corrected for multiple comparisons at the $p = 0.05$ level) accompanying the previously described two-way ANOVA confirmed that stereo motion judgments for light-from-above conditions were significantly different from the two other lighting directions for shadow speeds greater than approximately $0.2^\circ/s$ ($p$s < 0.016). No significant differences

Figure 2. The matched stereo motion speed plotted as a function of shadow speed. Data points represent the average of six observers and error bars signify 95% confidence intervals. In Figure 2A stereo motion and shadow speeds are plotted in degrees per second. In Figure 2B, both quantities are converted in centimeters per second. The dashed line in Figure 2B represents a perfect match between the 2-D shadow and stereo motion speed.
(at the 0.05 level) were observed between light-from-below and light-from-left conditions regardless of the shadow speed. This result is consistent with the original observations made by Kersten et al. (1996) and demonstrates the well described light-from-above prior that largely influences the perception of cast shadows (see Kleffner & Ramachandran, 1992; Ramachandran, 1988; Sun & Perona, 1998).

It has been proposed that cast shadows provide an indication of scene geometry as the 3-D location of objects and direction of illumination can be inferred from the orientation of a cast shadow. Kersten et al. (1997) demonstrated that, in the absence of additional depth cues, the position of a moving cast shadow relative to the object largely accounts for an object’s motion trajectory despite the fact that size of the stimulus might be inconsistent with the inferred motion direction. In the present study, we additionally find that the motion of the cast shadow largely accounts for the extent of illusory motion in depth for light-from-above conditions. This is evident when both the matched stereo motion and shadow speed are defined in equivalent values of centimeters per second. Note that in Figure 2A, stereo motion speed is given in terms of binocular disparity, and thus only describes the angular difference between the location of the stimulus on the retinas. However, geometric movement in depth of the stimulus (\( V^2 \), in centimeters per second) can be approximated from binocular disparity (\( \eta \) in radians) using the following equation: 

\[
V = D^2 \frac{\eta}{I}
\]

where \( D \) is the viewing distance (57 cm), \( \eta \) is the rate of change in disparity, and \( I \) is the interocular separation (6.2 cm).

Using this conversion, Figure 2A in terms of centimeters per second. A near perfect match (dashed line) between the speed of the shadow and stereo motion for light-from-above patterns is evident in this figure. Thus, in the absence of additional cues that may contribute to the geometry of the visual scene (such as diffuse shading or perspective), the visual system is highly dependent on information from cast shadows to interpret motion in depth. Additionally, this perceptual change is consistent with a scene geometry in which the light source is above and with its elevation 45° from horizontal. This configuration might reflect a default assumption made by the visual system to compute depth from shadows under the stimulus conditions employed in the present study. However, future empirical work is required to explore this possibility. Note that this is not the case for light from below and to the left patterns, but, rather, the same shadow speeds produce less observable illusory motion in depth effects. This finding indicates a lack of sensitivity to nonecologically valid patterns in the interpretation of cast shadows as has been suggested by a number of researchers (e.g., Khuu et al., 2012; Kleffner & Ramachandran, 1992; Rensink & Cavanagh, 2004). Alternatively, the visual system might interpret a different level of elevation for these two stimulus conditions. For example, given the physical speed of the shadow, the observed illusory motion for inverted shadows is consistent with a light source that is below and from the six o’clock position. Future studies investigating whether the visual system has a priori rules regarding the implicit position of the light source and cast shadows will be informative.

**Experiment 2: Illusory motion in depth and visual awareness**

In the previous experiment we showed that movement of a cast shadow could induce compelling illusory motion in depth of a stationary object. Moreover, the perceived motion is dependent on the lighting direction and is well accounted for by shadow speed for light-from-above configurations. In this experiment we examined whether the perception of illusory motion in depth requires the explicit awareness of the shadow or whether it occurs implicitly without awareness. Using the procedures of Experiment 1 in conjunction with a CFS paradigm, the moving cast shadow was binocularly suppressed while leaving the object visible. We questioned whether and how this affected the perception of the cast-shadow motion illusion. If the motion detection of the cast shadow reflects a low level and implicit process, we would expect the illusion of object motion to persist without awareness of the shadow. Alternatively, if the visual system were reliant on the cast shadow to explicate depth structure, then visual suppression of the cast shadow would likely attenuate the perception of illusory object motion.

**Methods**

**Observers**

Five experienced observers participated in Experiment 2. These were not the same observers as in Experiment 1. All had normal or corrected-to-normal visual acuity and were experienced in conducting psychophysical judgments of this nature.

**Stimuli and procedures**

The stimulus used in Experiment 2 was the same as in Experiment 1, but employed only the patterns implying light from above that had previously produced the strongest illusory effects. As in Experiment 1, a staircase procedure was used to match the perceived speed of two white squares whose motion was induced by a moving cast shadow or changing binocular...
disparity. This task was repeated for two experimental conditions in which the shadow was visible or suppressed from awareness using CFS (see Figure 3). For conditions in which the shadow was visible, as in Experiment 1, the white square and cast shadow were directly superimposed onto a dynamic Mondrian surface and both were presented to the left eye. In the CFS condition, the cast-shadow stimulus was presented to the left eye on a uniformly gray background (20 cd/m\(^2\)) and was made invisible by the presentation to the right eye of a dynamic Mondrian mask with its pattern randomized at 10 Hz but with a central square region (with blurred edges) left uniformly gray (see below for details). When viewed together, a stable and powerful binocular suppression was generated, with the Mondrian surface masking only the cast shadow. Because the white square overlapped with the central gray region of the mask it was the only feature of the cast-shadow stimulus visible under CFS conditions. This process of selectively suppressing components of the stimulus from awareness is similar to those adopted in previous CFS studies that have investigated how the perception of visual crowding and the formation of illusory contours (in Kanizsa figures) is dependent on the visual awareness of elements that induce these effects (see Harris et al., 2011; Ho & Cheung, 2011). Ten staircase trials were conducted for two stimulus conditions in which the shadow was visible or suppressed from awareness. Observers performed these 20 staircase trials in a randomized order.

Figure 3. How CFS was used to render the cast shadow invisible. Presented to the left eye was a moving cast shadow, while in the right eye a dynamic Mondrian mask was presented at 10 Hz. When viewed binocularly, the mask overlapped only with the cast shadow and the observers only saw the white square.
metrically blurred images separately to the eyes prevents binocular fusion (Meegan et al., 2001) and allows the image with sharper edges to dominate perception. In our procedure, this was vital to ensuring that binocular fusion and stereopsis did not cue the depth of the square, attenuating or abolishing the cast-shadow motion illusion. In a pilot experiment (which included the five observers who participated in the main experiment) we verified our methods by measuring the cast-shadow motion illusion (presented to the left eye) in the presence of a stationary blurred gray square (on a uniform gray background set to 20 cd/m²) presented to the right eye. For this we used a cast-shadow speed of 1.6°/s, (the speed that produced the greatest extent of illusory motion in Experiment 1), and conducted five staircase trials. The mean matched stereo motion speed for this condition was 0.11°/s (+ 0.03, 95% CL), which is similar to that reported in Experiment 1 for the same shadow speed. This finding indicated that the cast-shadow motion illusion persisted under these conditions (observers did not report seeing the gray region), which indicated that binocular fusion did not occur to ground the percept. Rather, the cast-shadow stimulus dominated perception. Thus, in the CFS conditions any reported attenuation of illusory motion in depth is not due to binocular fusion and the gray region of the mask cuing the position of the white square, but rather represents the masking of the cast shadow from visual awareness.

Prior to the experiment proper, we also quantified the effectiveness of the CFS mask in rendering the cast shadow invisible. In this pilot study, observers judged the position of the moving cast shadow (i.e., to the bottom right or bottom left of the object) while the cast shadow was suppressed from awareness using CFS. This judgment was repeated 20 times for each of the shadow speeds used in Experiment 1. The results of this control task are shown in Figure 4, which plots the proportion of times observers correctly identified the position of the cast shadow as a function of the cast-shadow speed. As shown, observers performed at the chance level over the range of cast-shadow speeds. These data indicate that, under CFS conditions, observers were not conscious of the cast shadow. Further, these results demonstrate that the range of speeds used in the present study did not break suppression to reveal the position of the cast shadow.

Results and discussion

The results of Experiment 2 are shown in Figure 5. This figure gives the average stereo motion speed required to match the cast-shadow stimulus for conditions in which the cast shadow was visible (squares) or invisible (circles) as a function of shadow speed. Error bars signify 95% confidence intervals. When the cast shadow was visible, increasing shadow speed increased the extent of illusory motion in depth.
This replicated Experiment 1. Linear regression analyses (as in Experiment 1) were performed on these data and the slope of the line of best fit (average $R^2$: 0.96, solid lines) was 0.068 ($\pm$ 0.098, 95% CI) and was significantly different from zero, $F(1, 28) = 182.6, p < 0.0001$. However, suppressing the cast shadow from awareness greatly reduced the perception of illusory object motion in depth (circles). The matched stereo motion speed for these conditions was approximately 0.01°/s, and this judgment did not change appreciably with increasing shadow speed. Moreover, the slope derived from the line of best fit was 0.003 ($\pm$ 0.098, 95% CI) and did not significantly deviate from zero, $F(1, 28) = 2.562, p = 0.1207$, indicating no effect of cast shadow speed. These results are consistent with the suggestion that the perception of depth from cast shadows largely requires the explicit representation of the position of the cast shadow and the object; without visual awareness of the cast shadow, the visual system is unable to effectively group it with the visible object to infer 3-D structure. Under these circumstances the visual system might default to other object properties such as the apparent size of the stimulus to cue the depth of the stimulus. Because the size of the object does not change across the stimulus presentation, this grounds the percept as largely stationary.

Importantly, CFS conditions presented the cast shadow stimulus to one eye and the dynamic Mondrian mask to the other. The cast-shadow stimulus was imaged onto a static and uniformly gray background. Note that in the shadow visible condition (also as in Experiment 1), the moving shadow was superimposed onto a dynamic mask, which were both presented monocularly to the left eye. In accounting for the results obtained from the CFS conditions in Experiment 2, it could be argued that the uniform background in the CFS conditions might produce less compelling illusory motion because the projection plane is not clearly referenced. In this case, the perception of illusory motion is substantially weaker because the cast shadow is not effectively grounded to the projection plane, with the visual system therefore unable to adequately compute the shadow-object separation to generate the percept of illusory motion. The dichoptically presented stimuli of Experiment 2 were viewed surrounded by a static square border and an identical Mondrian field, which are sufficient to reference the projection plane; observers informally reported compelling illusory motion with this stimulus configuration. However to confirm this, we repeated the cast-shadow visible condition, with the same observers, in an additional supplementary experiment that measured the illusory motion without a Mondrian background. In this experiment, the shadow stimulus was presented on a uniform gray background to the left eye, while no mask was presented to the right eye. When viewed binocularly, the percept was only of a square with its shadow cast on a uniform gray background.

The results of this additional supplementary experiment are shown in Figure 5 (triangles). These data show a systematic increase in the matched stereo motion speed with increasing shadow speed, much like when the cast shadow was projected onto a textured dynamic background (squares). Indeed, a repeated-measures two-way ANOVA (comparing the two shadow visible conditions across the different shadow speeds) reported a main effect of shadow speed, $F(4, 25) = 69.23, p < 0.0001$, but no difference between conditions in which the projection plane was textured or uniform, $F(1, 25) = 2.63, p = 0.1174$. Thus, we can suggest that the lack of illusory motion in CFS conditions is not due to a difference in backgrounds. This finding also suggests that the visual system inherently assumes a stationary projection plane in the perception of cast shadows and that the physical movement of the shadow is attributed to a foreground object moving in depth and not to a moving projection plane.

### Experiment 3: Cast-shadow motion adaptation and visual awareness

Experiment 2 demonstrated that when the moving cast shadow was suppressed from awareness by CFS, observers did not perceive the cast-shadow motion illusion. This finding suggests that the determination of depth from cast shadows is largely dependent on visual awareness. However, it is possible that the spatiotemporal properties of the dynamic CFS mask might not just attenuate visual awareness, but act to prevent the cast-shadow motion from being detected and processed by the visual system in the first place. Thus, in under CFS conditions in Experiment 2, the cast-shadow motion is not available for processing as it is simply destroyed by the mask. To rule out this possibility in Experiment 3 we conducted a motion adaptation experiment to determine whether the cast-shadow motion is “unconsciously” detected by the visual system under CFS conditions.

### Methods

**Observers**

Four of the five observers who participated in Experiment 2 acted as observers in this experiment. All were naive to the aims of the experiment but were experienced psychophysical observers.
Stimulus and procedures

Previous studies using binocular rivalry and CFS have well demonstrated that a motion after effect (MAE) can be produced without awareness of the adapting stimulus (e.g., Kaunitz, Frascasso, & Melcher, 2011; Lehmkule & Fox, 1975; Maruya, Watanabe, & Watanabe, 2008; O’Shea & Crassini, 1981). In Experiment 3 we adopt this adaptation paradigm to determine whether the cast-shadow motion stimulus survives CFS suppression. In particular the cast-shadow stimulus used was a black square moving at one speed of 1.6°/s and in a direction consistent with a light-from-above configuration. As in Experiments 1 and 2, a white square (representing the object) was presented and visible to the observer. There were three experimental conditions. In Condition 1 the moving cast shadow was visible and superimposed over a dynamic Mondrian background (as in Experiment 1), which was presented only to the left eye. While maintaining fixation at the center of the screen, the cast shadow was presented initially 4° to the right of fixation and moved down and to the right in a 15° direction. The cast shadow traversed a total distance of 3.2° along this trajectory before being “wrapped” around to its original starting position. As the stimulus became increasingly blurred (to simulate the shadow preumbra) this minimized the abrupt disappearance as it was replotted back to its starting position.

To measure the strength of the MAE in the abovementioned conditions, we examined the effect of the MAE on the perceived position of a line placed within the region of adaptation. Previous research (e.g., Nishida & Johnston, 1999; Snowden, 1998) has well demonstrated that motion adaptation, in addition to generating an illusion of motion, also distorts the perceived position of objects in the direction of the MAE. Accordingly, quantifying whether adaptation to the cast-shadow motion under CFS conditions distorts the position of a line provides an objective means of determining whether the cast-shadow motion is detected by the visual system and not simply negated by the CFS mask. We used a Vernier acuity task in which observers had to judge the position of a line presented to the adapted area relative to another line unaffected by adaptation. In particular, after an initial period of adaptation of 20 s, the stimulus was removed from the screen and two horizontal black lines (12 cd/m², 0.25° × 0.75°) were briefly presented (187 ms) 1.25° directly below the point of fixation. These lines were horizontally separate by 0.25° and positioned on either side of the midline of the screen. This ensured that the line on the right fell within the adapted region. The physical position of the right line was fixed and an Method of Constant Stimuli (MoCs) was used to physically offset the left line vertically at levels of −0.4°, −0.3°, −0.2°, −0.1°, 0, 0.1°, 0.2°, 0.3°, 0.4°, and 0.5° relative to the position of the right line. Positive values indicate that the left line was physically above the right line, while negative values indicate that it was below. Each of the 10 spatial positions was repeated 20 times in a randomized order with top adaptation periods of 4 s between presentations. On each trial, the observers were required to indicate whether the left line was above or below the right line.

We repeated the abovementioned procedure in two additional conditions, but with the cast shadow masked using CFS (as per the methods described in Experiment 2). This CFS condition was conducted with the cast shadow moving (Condition 2) and with it stationary (Condition 3 and presented at a random position along the trajectory of motion). The latter condition provided a baseline comparison, as it is expected there will be no MAE. In relation to the objectives of Experiment 3, if dynamic CFS prevented the processing of the cast-shadow motion an expectation is that there will be no MAE, and therefore, the perceived position of the line will not be distorted. Alternatively, if the CFS mask works to suppress the cast shadow from awareness but does not prevent its processing (it remains to be unconsiously detected), then it is expected that motion adaptation will occur and this will distort the position of the line in the MAE direction.

Results and discussion

The results of Experiment 3 are shown in Figure 6A. The proportion of times in which the left line was judged to be above to the right line is plotted as a function of the physical separation between the two lines. The results for the three different stimulus conditions are shown as different symbols, error bars represent 95% confidence intervals. Logistic functions fit to data (using Graphpad Prism 6) are also shown in Figure 6A (different gray scale lines, average $R^2$: 0.94), and they were used to estimate of the point of subjective alignment (PSA), which represents the physical position of the left line that is perceptually aligned with the right line. PSA values are shown as bar graphs in Figure 6B with error bars signifying 95% confidence intervals. A number of findings are evident. First, when the cast-shadow motion was visible (triangles in Figure 6A) a position shift was observed. This indicated that the cast-shadow motion was sufficient to generate an MAE, which distorted the position of the right line. Note in Figure 6B the PSA for this condition indicated that the left line was offset by approximately 0.19° in the MAE direction for it to be perceptually aligned with the left line. When the cast shadow was stationary and masked from awareness (circles in Figure 6A) no MAE was produced and the PSA for this condition was approximately zero.
However, when the cast shadow was moving (squares in Figure 6A), the PSA was approximately 0.14°, which indicated a position shift in the direction of the MAE despite the observer not being aware of the cast-shadow motion.

A repeated-measures one-way ANOVA indicated that the mean PSAs for the three stimulus conditions were significantly different, \( F(2, 3) = 52.94, p = 0.003 \). Post-hoc comparisons (Tukey corrected for multiple comparisons at an alpha of 0.05) indicated that there was a significant difference between the PSA for the condition in which the cast shadow was stationary and when it was moving (visible: \( p < 0.0012 \); invisible: \( p < 0.006 \)) but no difference between visible and invisible moving cast-shadow conditions (\( p = 0.275 \)).

The results of Experiment 3 demonstrate that the CFS mask used in the present study acted primarily to suppress the cast-shadow motion from visual awareness and did not prevent from being processed by the visual system. We note that an MAE was produced under CFS conditions as evidenced by a shift in the apparent position of the line placed within the adapted region. The findings of Experiment 3 are entirely consistent and replicate previous studies that have demonstrated that CFS does not directly prevent the processing of motion but rather modulates the visual awareness of it. For example, using CFS to suppress the adapting stimulus from awareness, Maruya et al. (2008) observed MAEs for both static and dynamic test patterns presented to the adapting eye as well as to the other eye, which demonstrates interocular transfer of the effect. Additionally, Kaunitz et al. (2011) using CFS have reported that motion adaptation to invisible stimuli is also observed with complex global-motion stimuli (e.g., spiral), which is believed to be processed at higher stages of visual processing. In conclusion, Experiment 3 confirmed that the cast-shadow motion is unconsciously detected by the visual system, but visual awareness is necessary for it to be used as a depth cue to infer the motion in depth of objects.

**General discussion**

In this study we examined the role of visual awareness in the perception of depth from cast shadows. We did so using the cast-shadow motion illusion, originally reported by Kersten et al. (1996), because it is a powerful example of how the visual system actively infers the 3-D configuration of objects from cast shadows. In Experiment 1, we quantified the cast-shadow motion illusion in a matching experiment: The perception of illusory object motion was matched to another square undergoing stereo motion. Results indicated that, as cast-shadow speed increases, the
stereo motion speed required to match the illusory motion also increases. However, the extent of this effect was dependent on lighting direction, with light-from-above conditions producing the strongest illusory motion effect. Moreover, for light-from-above conditions, the shadow speed accounts well for the extent of illusory motion in depth. It was noted that a change in shadow speed produced an equal change in illusory object motion in depth. Using CFS in Experiment 2, we showed that suppressing the cast shadow from visual awareness greatly attenuates the perception of illusory motion in depth; matched stereo motion speed was close to 0°/s. Further, changing shadow speed—which increases the perception of illusory object motion when the shadow is visible—had no effect when the cast shadow was invisible under CFS conditions. In Experiment 3 we ruled out the possibility that the elimination of illusory object motion in depth is because the CFS masks prevent the processing of the cast-shadow motion. We demonstrated that the cast shadow continues to be unconsciously processed by the visual system sufficient to generate an MAE, which distorts the position of a briefly presented line. In summary, the results of this study suggest that the visual system is able to quantitatively derive an estimate of depth and position from moving cast shadows, but this process is largely dependent on visual awareness. Whether this is also the case for other monocular cues to depth structure is beyond the scope of the present study but is a promising avenue for future research.

As noted in the Introduction, a number of studies have indicated that cast shadows are rapidly detected and reflect processing in early cortical stages (e.g., Castiello et al., 2003; Dee & Santos, 2011; Rensink & Cavanagh, 2004). Contrary to these, we find that visual awareness is necessary for the perception of illusory motion in depth induced by a moving cast shadow. This distinction is consistent with the assumptions made in the Introduction that the processing of cast shadows might reflect two functionally distinct stages. First, cast shadows are quickly and implicitly interpreted (involving a coarse scale analysis of the visual scene) by the visual system. This information contributes to the identification and the recognition of objects. Second, the position of the cast shadow (relative to the object) is used infer the object’s 3-D position in the scene. The present study suggests this latter process is likely driven by visual awareness, which might reflect an additional step in which the cast shadow and the object must be explicitly coded and perceptually grouped to determine their spatial separation. The distinction between the present and previous findings can be understood in terms of the two-stage model of processing: We suggest that previous studies have only considered the first stage involved in shadow processing. For example Castiello et al. (2003) noted that patients with visual neglect remain sensitive to cast-shadow information that facilitated their recognition of objects. This only reflects low-level processing relating to the initial identification of cast shadows. However, the explication of depth from cast shadows is necessarily dependent on visual awareness of cast shadows, which occurs later in processing after shadows have been identified. Future studies examining the perception of cast shadows are needed to explore fully the distinction between the detection of cast shadows and how they are used as a cue to determining the depth position of objects.

Additional evidence that supports the notion that the extraction of 3-D structure from moving cast shadows might reflect a two-stage process comes from previous binocular rivalry and CFS studies that examined the role of visual awareness in processing of local and global information. It has been argued that while the detection and processing of low-level visual features such as local orientation, form, brightness, and unidirectional motion can occur under binocular suppression (see Blake & Fox, 1974; Lehmkuhle & Fox, 1975; Maruya et al., 2008; Wade & Wenderoth, 1978), the analysis of high-level features are not processed when suppressed from awareness (see Blake, 1997; Lin & He, 2009). For example, previous studies (see Moradi, Koch, & Shimojo, 2005; Stein & Sterzer, 2011) have well established that the processing of faces is largely dependent on visual awareness. In particular, face configural aftereffects are only obtained under conditions in which the observer is aware of the adapting face. Thus CFS might be more effective in disrupting perceptual grouping and integration of local information in the detection of both faces and cast shadows. There are exceptions (e.g., emotional words and manipulable tools), though these appear to be observed by specialized pathways in the visual processing hierarchy (Almeida, Mahon, Nakayama, & Caramazza, 2008; Tsuchiya & Koch, 2005).

As noted in Experiment 3, the lack of illusory motion reported in the present study is not because CFS prevents the processing of local motion; previous studies have shown the processing of local motion can survive binocular suppression (see Kaunitz et al., 2011; Lehmkuhle & Fox, 1975; O'Shea & Crassini, 1981). Rather, the lack of illusory motion is due to a failure in processing 3-D structure inferred from the position of the moving cast shadow. The fact that CFS disrupts the perception of illusory motion from cast shadows suggests that this process might occur at higher levels, and the perceptual pairing of the moving cast shadow and its casting object requires conscious awareness. Similar effects have been reported by Harris et al. (2011), who observed that while illusory brightness of a
target flanked by a different contrast surround persists under CFS conditions, illusory contours in Kanisza figures (produced from the perceptual grouping of discrete tokens that indicate the corners of the stimulus) do not. Arguably, the former reflects processing at the receptive field level and the latter is dependent on perceptual grouping, which might require awareness. Additionally, neural imaging evidence exists demonstrating that the ventral stream of visual processing, which is responsible for the analysis of form information, is not activated when the stimulus is suppressed from awareness using CFS (see Fang & He, 2005). Because cast shadows are used by the visual system as a form cue to the object’s 3-D position, this neural evidence is consistent with the dependency of shadow processing on awareness.

In conclusion, the results of the present study suggest that the visual system is able to extract an estimate of metric depth from a moving cast shadow, with the strongest effects selective for configurations in which light is from above. However, while the initial detection of cast shadows might not be driven by conscious perception, it appears that the inference of 3-D structure is. The explication of depth structure most likely occurs in higher cortical stages whose processing is affected by CFS.

Keywords: cast shadows, depth perception, visual awareness

Acknowledgments

We thank the observers who participated in the study and Joanna Kidd and the two anonymous reviewers for their helpful comments and suggestions. This research was supported by an Australian Research Council (ARC) Discovery Project Grant (DP110104713) to S. Khuu.

Commercial relationships: none.
Corresponding author: Sieu Khuu.
Email: s.khuu@unsw.edu.au.
Address: The School of Optometry and Vision Science, The University of New South Wales, Sydney, NSW, Australia.

References


following adaptation to visual motion. *Current Biology, 8,* 1343–1345.


