Perception, 2004, volume 33, pages 1369 – 1383

DOI:10.1068/p5320

Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art

Jayme Jacobson, Steffen Werner
Department of Psychology and Communication Studies, University of Idaho, Moscow, ID 83844-3043, USA; e-mail: jaymejacobson@moscow.com
Received 2 February 2004, in revised form 20 April 2004; published online on 8 December 2004

Abstract. The kinds of visual cues artists choose to use or not use in their work can offer insight into perceptual processes. On the basis of the observed paucity of the use of cast shadow in pictorial art, we hypothesized that cast shadows might be relatively expendable as pictorial cues. In this study, we investigated two potential reasons for this expendability: first, viewers might be insensitive to much of the information that cast shadows provide; and, second, ambiguities about what is shadow and what is pigment can often be resolved only through motion—something that static media are ill-equipped to deal with. In experiment 1, we used a visual-search paradigm in which viewers had to determine if there were odd cast shadows in sets of 4, 8, 16, and 32 objects. In experiment 2, viewers had to discriminate between shadow/pigment ambiguities in both still and moving images. Our results demonstrate that viewers are neither particularly sensitive to static cast-shadow incongruities, nor are they able to disambiguate cast shadow from pigment without continuous motion information. Taken together, these results may help explain why cast shadows are relatively rare in static pictorial work.

1 Introduction

One of the most overlooked sources for information about human perceptual processes is the vast body of work produced by professional artists and artisans throughout history, from the first attempts at pictorial representation, through all of the varying styles and techniques of the intervening millennia, to the art of the present day. This oversight is unfortunate because these art products constitute what is, arguably, the largest archival data bank of human perception available to researchers. Whereas the art of children, the art of schizophrenics, and the art of many other specialized communities has always maintained a respected place within the tradition of perceptual investigation, art that hangs in gilt frames, that is housed in museums, and speculated upon by informed critics has been considered largely off-limits to researchers: too bounded by the aesthetic conventions of the various eras that produced it to yield untainted information about human perception in general. Of course, there are exceptions to this rule. Cave paintings, some of the first known representational artifacts, have stimulated research interest in anthropology and evolutionary psychology alike (Lewis-William 2002), and recently psychological and neuroscientific approaches to the study of art, aesthetic experiences, and selective artists have become more frequent and popular (eg Ramachandran and Hirstein 1999; Zeki 2001). But approached as an enormous archival repository—with large sample sizes of broad cross-sections of pictorial art in order to reduce variability due to idiosyncracies of individual artists, cultural aesthetics, and style trends—mainstream art can, at the very least, give researchers broad hints about the ways in which visual information has been perceived across the ages.

This fossil record of perception left by artists through their works is of interest to the perceptual researcher not because artists are more perceptive than non-artists, but because of the multi-step process that artists must go through to produce their work.

Like anyone, artists must first take in sensory information, assess it, and transform it through mental processes. Then, what has been judged relevant information must be routed back out onto the canvas or other media. At this point, because artists are, of necessity, always their own first audiences, they must re-perceive and evaluate the newly laid down information. This evaluation takes into consideration not only the real world as it is sensed, but also an internal representation that the artist has formed from those sensations. After the evaluation, the work is adjusted to bring it into alignment with this internal representation, and the cycle of re-perception, evaluation, and adjustment continues until a satisfactory 'fit' is achieved. In this sense, "visual art ... obeys the laws of the visual brain, and thus reveals these laws to us" (Zeki 2001, page 52).

During the perceptual and representational cycle, certain perceptual cues will be accorded more significance than others, and some sensory information that is present in the visual scene will be considered sufficiently insignificant—or even distracting enough—to be left out of the work altogether. Thus, for the production of artwork, perceptual cues can be divided into two broad categories: cues that are essential, and cues that are expendable. Expendable cues are those that are either of negligible benefit to the work as a whole, or those that have been considered too distracting with regard to the pictorial representation. For the researcher, expendable cues can offer as much insight into perceptual processes as essential ones.

Returning to the idea of artwork as an immense archival database, one pattern that seems to emerge from the body of Western Art is the relative absence of *cast shadows*: shadows that are cast onto other surfaces. In contrast, *attached shadows*, which have been used in the form of shading and modeling to give depth and form to objects, have been used extensively throughout the history of art. This situation became fairly evident to the curators of the National Gallery in London when they attempted to compile enough work for the 1995 exhibition, *Shadows: The Depiction of Cast Shadows in Western Art*. In fact, the accompanying catalogue for the exhibition noted the surprising paucity of cast-shadow depiction available from such an extensive collection of pictorial work (Gombrich 1995).

Are cast shadows used less frequently than attached shadows in works of art? We decided that the frequency of cast-shadow depiction would make a good candidate for a small straightforward archival analysis. Using *Janson's Story of Painting* (Janson and Janson 1984), we asked six undergraduate volunteers from the University of Idaho (four of them male), to evaluate 89 pictorial works for shadow usage. The art works were numbered in the book and participants were asked to fill out a corresponding numbered form indicating whether or not a given work made use of shading (attached shadows), grounding shadows (the unspecific shadow cast by an object directly touching the ground beneath it), or cast shadows. The results were surprisingly unreliable. Not only were the inter-rater correlations unacceptably low (between $-0.06 < r < 0.52$ for shading, $-0.06 < r < 0.39$ for grounding shadows, $0.12 < r < 0.46$ for cast shadows), there did not seem to be a clear and shared understanding about how to interpret the shadows, whether the shadows were consistent across a scene (eg in the same direction), and whether shadows were cast by only a few or a majority of the objects in the scene. An informal comparison of ratings produced by art professionals showed a similar problem. The only consistent result across raters was that the attached shadows (mean = 80%) were much more frequent in the collection of images sampled from the textbook than both grounding shadows (mean = 54%) and cast shadows (mean = 53%).

It would appear, then, that cast shadows are used less frequently in pictorial works than shading, even though cast shadows are just as prevalent in real life. Is it possible that cast shadows go largely unnoticed during the production of pictorial work? Gombrich (1995) cautioned against any assumption that the absence of cast shadows indicates inattentiveness on the part of the artist. But if artists are fully aware of the cast shadows in the scenes they are attempting to reproduce, then the infrequency of cast-shadow inclusion is a signal to researchers that cast shadows are somehow winding up in the expendable-cues category. The question is, why?

One explanation may be that the theory and study of the geometric projection of shadows, skiagraphy, is extremely complex. The problem with this explanation is that, in practice, few artists would attempt to calculate the precise geometries of cast shadow in a scene when convincing results could be obtained, with much less effort, by clear, concentrated observation. Casati (2003) suggested that the reason for the relative absence of cast shadow in pictorial work might be a very simple one: the subtleties of cast shadow are difficult to render well. Again, this is doubtful. Artists regularly tackle difficult and subtle subject matter. While the handling of cast shadow requires finesse, it arguably takes more skill to render water in all of its moods; the various textures of velvet, silk, and fur; and the reflective qualities of glass, than to render cast shadows.

We would like to suggest an alternative hypothesis for the paucity of cast shadows in pictorial work, one that is two-fold and considers how expendable cast shadows are as pictorial cues. We earlier defined expendable cues as those that are either of negligible benefit to the work as a whole, or those that have been considered too distracting with regard to the representation. First, we postulate that incongruities in static shadow representations may go largely undetected by the viewer, meaning that the exclusion of cast shadow would have little consequence on the viewer's ability to interpret a scene. Second, cast shadows may convey visual information that pictorial work is poorly equipped to deal with: information about transience and movement. When analyzed with respect to static, pictorial representation, cast shadows might even have the potential to create distracting ambiguities that run counter to the intentions of the artist.

The first part of our hypothesis seems odd at first—why would the visual system, which has evolved under natural lighting conditions with ample opportunity to view objects and their cast shadows, not be sensitive to singularities in the otherwise predictable and regular distribution of cast shadows? After all, cast shadows can provide important information about the shape of an object—especially its location with respect to other objects and the ground plane [see Kersten et al (1997) for a striking example] and provide a statistical invariant in the visual environment. However, in a simple change-blindness study looking at the role of expertise in detecting semantic and non-semantic changes in visual scenes (Werner and Thies 2000), the second author found that one of the hardest changes to detect was whether a non-semantic change was introduced by changing the orientation of a single cast shadow. Even though this change led to an inconsistent scene, observers did not seem to pay attention to the shadow orientation. In a related study, Ostrovsky et al (2001) showed that observers do not automatically recognize illumination inconsistencies in displays of multiple, randomly oriented cubes. In their study, the shading of one of the cubes was inconsistent with the light source that was used to illuminate the remaining cubes. That study, similar to both the classic study by Enns and Rensink (1990) and our own, used a visual-search paradigm to show the inefficiency of searching for the odd item. Their findings show surprising difficulties in identifying illumination inconsistencies compared with Enns and Rensink's original finding of efficient search for some 3-D displays. Interestingly, their paper at first glance seems to include the detection of odd cast shadows in regular arrangements of homogenous objects in the 'perceptually easy' category—unlike the results we will present later in this paper with exactly these types of displays.

The second part of our hypothesis deals with cast shadows as conveyors of transitional or temporal information, along with their ambiguous, and often distracting nature in static displays. Cast shadows can be used to convey temporal information and are frequently seen in works that attempt to evoke a certain mood based on the time of day or season. In fact, Giorgio de Chirico often paired bright, noonday skies with long afternoon shadows to create a sense of arrested time (Soby 1969) (see figure 1). But it is this temporally based, transient nature that can make cast shadows a liability in pictorial works. Cavanagh and Leclerc (1989) observed that there is often a potential for confusion between pigment and shadow: that in real life, the visual system must discriminate between the two. Normally, this discrimination takes into consideration what is most likely, based on various cues such as relative positions, tonal values, etc. But sometimes these cues are ambiguous, making it difficult to determine what is what. In a real-world situation, this ambiguity can be resolved through movement: either movement by the objects in the scene, movement of the objects that cast shadows, or movement of the light source. This resolution is due to the transient nature of shadows: in movement, they interact differently with the surfaces of a scene than pigment does. However, in pictorial works, this disambiguation is not possible. The pictorial world has no recourse to movement: neither objects nor light sources in the scenes of a static work can move.

Figure 1. Giorgio de Chirico's *Delights of the Poet*, 1913. De Chirico often paired bright, noonday skies with long shadows to create a strange sense of timelessness.

In the pictorial world, the essential problem specific to cast shadows is that, by their very definition, they must be cast onto another surface. If that surface is relatively free of critical information, such as in the case of open ground, then there is no conflict, no serious potential for confusion, or ambiguity. But in the real world, cast shadows fall on all kinds of inconvenient surfaces. In general, artists take care to arrange their scenes so that critical information will not be obscured by other, less vital, elements. Inconvenient shadows can be controlled by careful lighting, by arranging objects so that there is no obscuration, or by simply excluding the more problematic cast shadows. An example of this can be found in the frescos at the Brancacci Chapel. Casati (2003) has observed that many of the shadows in the chapel's frescos are oddly transparent.

However, one of the frescos, Masaccio's *St Peter Healing with His Shadow*, nicely illustrates the problem that cast shadows can create in static media and offers a clue as to why the shadows might be transparent (see figure 2). The scene is shown at the moment that Saint Peter's healing shadow is passing over the foremost figure. Both the saint's shadow and the expression on the figure's face are critical to the narrative. This leaves the artist with the problem of how to depict the shadow without obscuring the face. Masaccio appears to have resolved this dilemma by making the shadow visible on the ground, but transparent over the figure.

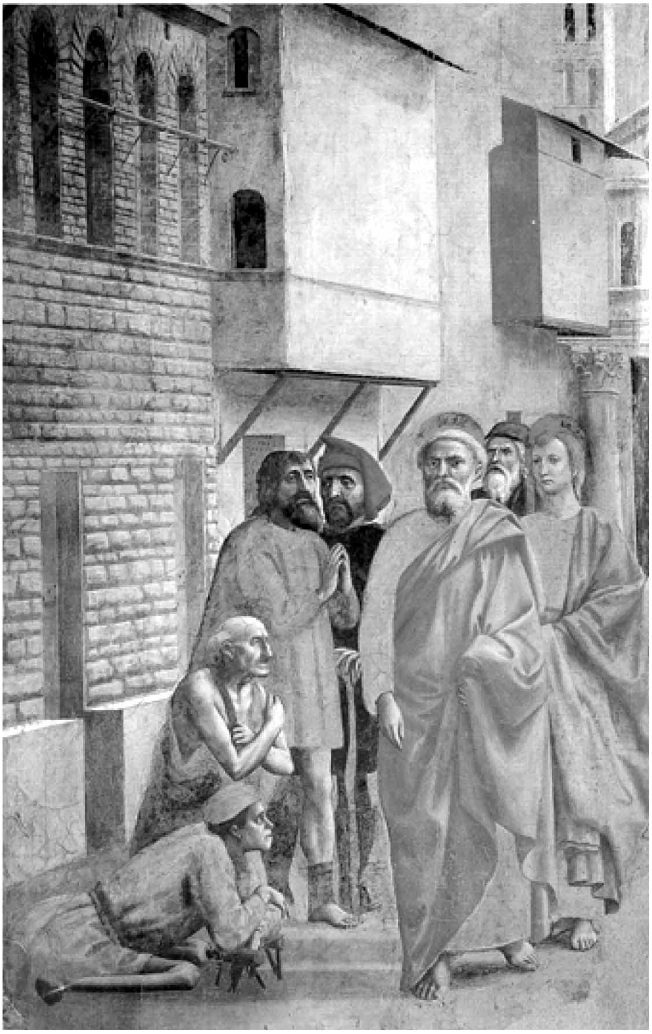


Figure 2. Masaccio's *St Peter Healing with His Shadow*, 1424 – 1428. By using transparent shadows, Masaccio avoids obscuring critical elements in the fresco.

To overlay a figure with the shape(s) of cast shadow is to obscure it, to degrade its unity, to camouflage it. Camouflage is generally antithetical to the goal of a pictorial work, which seeks to use readable elements to convey meaning. There have been exceptions to this, of course—particularly in modern art. Behrens (1998) has pointed out the relationship between camouflage and Cubism but, historically, the goals of representational art and camouflage have been at odds with each other.

Writing about the natural kinds of camouflage found throughout the animal kingdom, Cott (1966) observed that "nature goes to work in the opposite way to an artist—who by use of light and dark pigments produces the illusion of solidity and relief on a flat surface" (page 124). Natural camouflage, on the other hand, aims to obscure, flatten, and de-solidify the form of an animal—to hide it from both predator and prey.

Thayer (1909) observed that many animals utilize markings that aid in concealment by visually breaking up the form of the animal. This *obliterative coloration* can also mimic cast shadow. Even in man-made camouflage, the effect of cast shadow has always been an important consideration (Luckiesh 1965). In fact, the incorporation of cast-shadow tones into existing camouflage patterns can dramatically enhance concealment potential (Mossy Oak 2003). But for all forms of camouflage, the crucial condition is stillness. The instant that there is movement on the part of the camouflaged element, the effectiveness of any form of concealment is immediately diminished. Due to the Gestalt law of common fate, elements moving in the same direction are perceived as being united in form.

The second experiment that we report in this paper deals with this aspect of our hypothesis. The inclusion of cast shadows in a piece of pictorial art can lead to ambiguity about what is pigment versus what is shadow. In experiment 2, we show how the static view of an object, in which pigment and cast shadow are ambiguous, cannot be resolved by the visual system. In contrast, increased amounts of continuous movement information lead to increases in the ability to distinguish pigment from shadow. We also look at the role of alternative pictorial cues, like the penumbra of a shadow, and its role in signaling the presence of shadow in static media.

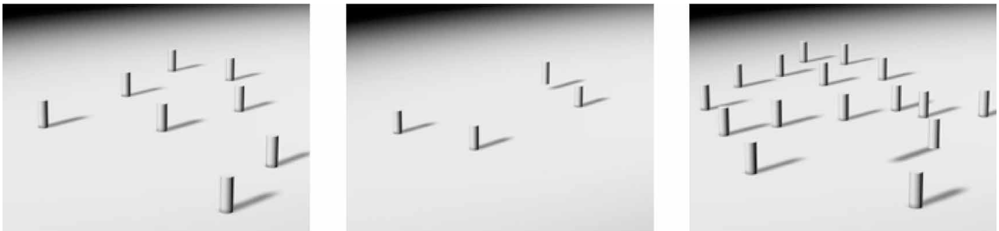
2 Experiment 1

The first experiment addresses the first part of our hypothesis: that human observers are insensitive to local inconsistencies of cast shadows. Like the Ostrovsky et al (2001) study, which we only discovered after we had run part of our experiment, we decided to test the observers' ability to find inconsistencies in cast shadows with a simple visual-search paradigm. To rule out that the complexity of the scene would make it hard for observers to find inconsistent or odd cast shadows (eg by cluttering the display with irrelevant background information), we chose to test our hypothesis by using simple, three-dimensional scenes with only one type of object on an homogenous background.

2.1 *Methods*

2.1.1 *Participants.* Twenty undergraduate and graduate students from the University of Idaho (eleven of them male) participated in the study (mean age = 27.7 years, SD = 12 years). Most of them participated for course credit. All of the participants had normal or corrected-to-normal vision.

2.1.2 *Materials.* We constructed two sets of 160 gray-scale images (800×600 pixel resolution, Maxon Cinema 4D R8 renderer) that showed arrays of 4, 8, 16, or 32 gray cylinders on an homogenous gray plane. The odd cast shadow for the first set of images differed in orientation: half of the scenes per condition were homogenously lit by one light source, which produced a cast shadow for each object on the ground plane. Each object was also lit from a second light source from the opposite side (180° orientation change, same elevation and distance). This second light source did not cast a shadow. For the other half of scenes per condition, the first light source did not produce a cast shadow for one object. In these cases, the second light source provided the illumination for that object's shadow, resulting in a cast shadow that ran exactly opposite to the direction of the other cast shadows. We chose a difference in orientation by 180° to avoid simple image-matching strategies that would allow observers to identify the odd shadow solely on the basis of orientation (see figure 3 for an example).



(a) same (b) odd elevation (c) odd orientation

Figure 3. Examples of the stimuli used in experiment 1. (a) No target present; (b) odd elevation (producing a separation between the object and its cast shadow); (c) odd orientation of one cast shadow.

The second set of 160 images used a different principle to produce the odd shadow. Half of these images were again produced in the same way as before, resulting in a homogenous set of cast shadows. However, to introduce an odd shadow in the other half of the images, one of the cylinders was elevated by 50% of its height from the ground plane, thus producing a separation between the shadow and the object. In the animations used by Kersten et al (1997), this was a very powerful cue for vertical height. Unlike the first set of images, the odd shadow therefore differed not in orientation but in its spatial separation from the object.

For each of the two image sets, we produced an equal number (40) of images for each set size (4, 8, 16, and 32), half of which were homogenous, and half of which included an odd cast shadow. To produce the images, we used two different camera elevations and opposite camera positions within the scene, resulting in four different views of each scene.

2.1.3 *Procedure.* The experiment was run with the experimentation software Presentation 0.76 (http://www.neurobs.com) to control the display and timing of stimuli and response measurement. The stimuli were presented on a 17 inch CRT display (1024 $\times$ 768 pixels, 75 Hz). Participants were instructed to press the 'D' key for 'odd' trials and the 'K' key for 'same' trials. They were asked to respond as quickly as possible while keeping their error rate to a minimum. To reduce error, an error feedback sound was presented in case of a misclassification. The type of change (orientation versus elevation change) was blocked and the order balanced across participants. Within each set of 160 trials, the order of trials was randomized.

Participants initiated each trial by pressing the space bar. Immediately after the key-press, an image appeared that had the same dimensions as the test stimulus image. For the elevation condition, we used a homogenously gray background. In the orientation condition, each test stimulus was preceded by its appropriately shaded background to aid the participant's understanding of the lighting conditions in the subsequent test image. After 500 ms, the background image was replaced by the test stimulus. The test stimulus stayed on the screen until the participants indicated their responses. Between conditions, participants were instructed to change their search from either the orientation condition to the elevation condition or vice versa.

2.2 *Results*

In the odd-orientation condition, participants correctly identified the odd shadow 87% of the time (hits) and falsely reported the presence of an odd shadow 2% of the time for target-absent (same) trials (false alarms). The average response time was 1888 ms (SD = 420 ms) for correct odd trials and 2850 ms (SD = 797 ms) for same trials. In both cases, the average response times increased monotonically with increasing set size from 1141 ms (odd trials) and 1267 ms (same trials) for set size 4, to 3148 ms (odd trials) and 5139 ms (same trials) for set size 32 (see figure 4).

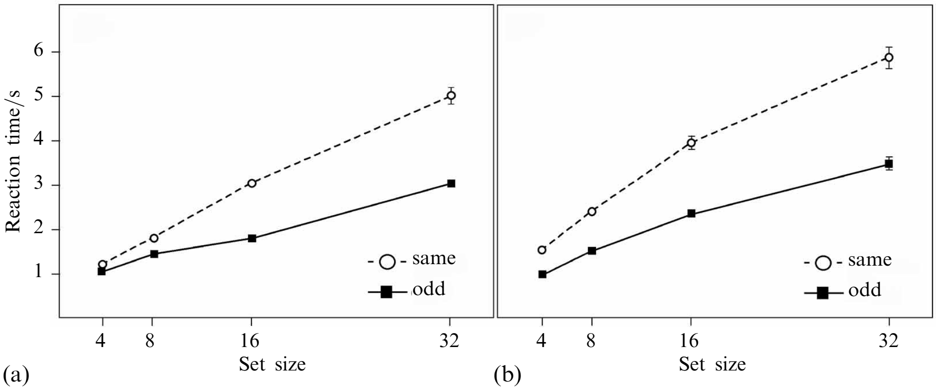


Figure 4. Mean reaction times for correct responses for target-present (odd) and target-absent (same) trials. Error bars indicate standard errors. (a) Orientation condition; (b) elevation condition.

The results for the elevation-change condition were similar. Participants had an average of 84% hits for the odd shadow and a rate of 3% of false alarms. The average response times were again shorter for odd trials (mean = 2150 ms, SD = 592 ms) than for same trials (mean = 3480 ms, SD = 1149 ms). As in the orientation condition, response times increased with increasing set size. For the odd trials, they ranged from 1072 ms for set size of 4, to 3531 ms for set size of 32, while the same trials saw even a larger increase, 1575 ms (set of 4) to 5925 ms (set of 32).

The average response times of each participant in each of the 16 conditions were submitted to a three-factorial, repeated-measures ANOVA with the within-subject factors shadow manipulation (orientation versus elevation), set size (4, 8, 16, and 32), and target presence (odd shadow versus all shadows same). Because of significant non-sphericity of the data, the degrees of freedom were Greenhouse–Geisser corrected. As should be evident from figure 4, all three main effects were significant. Reaction times were significantly faster in the orientation condition than in the elevation condition ($F_{1,19} = 14.6$, $p = 0.001$), reaction times increased with set size ($F_{1.1,20.5} = 98.1$, $p < 0.001$), and reaction times were significantly faster when an odd shadow was present ($F_{1,19} = 90.7$, $p < 0.001$). In addition, all second-level interactions were significant. Reaction times increased more with increasing set size in the elevation condition than in the orientation condition ($F_{1.3,24.5} = 8.0$, $p = 0.006$), and the effect of set size was more pronounced for target-absent (same) trials ($F_{1.3,25.5} = 51.4$, $p < 0.001$). Similarly, the difference between elevation and orientation conditions was larger for target-absent (same) trials ($F_{1,19} = 11.2$, $p = 0.003$). The remaining interaction between all three factors was not significant ($F_{1.6,30.7} < 1$).

The increase of response time (RT) with set size largely followed a linear trend. For all four conditions (orientation versus elevation × target present versus target absent), the product-moment correlations between set size and RT were $r > 0.99$. In the orientation condition, the slope of the regression function was 76 ms per object for target present and 139 ms per object when all the shadows were homogenous. In the elevation condition, the respective slopes were 95 ms per object versus 157 ms per object. To assess the consistency of the RT increase with increasing set size, we looked at each individual's pattern of result. Of the 240 pairs of neighboring means (4×3 per participant), only 5 pairs (2%) involving three separate participants did not follow the predicted strict monotonic trend.

2.3 *Discussion*

The results of experiment 1 show a clear picture. When the observer was looking for an inconsistent (odd) cast shadow among a set of homogenous shadows, his/her search time increased monotonically with increasing number of relevant objects (distractors) in the display. It appears that there is a constant additional time interval of approximately 140 ms required for the processing of each additional item in a display—dramatically more time than the estimated 6 ms for the detection of lighting inconsistencies for shaded cubes reported by Enns and Rensink (1990, their experiment 1A) as an example for efficient search. These results hold for both types of odd shadows investigated in this study. The slopes of the linear search functions ranged between 76 ms per item to 157 ms per item, which indicates that visual search for inconsistencies in cast shadows is not efficient and probably relies on high-level controlled processes that work in a serial fashion. It therefore corroborates the findings of Ostrovsky et al (2001) that inconsistent shading of objects is difficult to detect. Even though we found a difference between the two types of inconsistencies that we introduced in this experiment, we do not feel justified in generalizing this result to the ease of identifying odd cast shadows by orientation versus identification by spatial separation. Different parametric values for the chosen variables (eg a larger spatial separation or a nonplanar ground plane) might lead to different magnitudes of the effects.

Besides the general increase of RT with set size, the study further shows that a lack of sensitivity for inconsistent illumination exists even in scenes that are simple and highly regular. We would predict that the effect is even larger for cluttered displays with multiple object shapes and nonplanar surfaces. In light of our main research question, these results strongly underline the insensitivity of the human visual system to inconsistencies in the arrangement, or even presence, of cast shadows. Even though cast shadows do convey important information about the arrangement and spatial position of objects in the scene (eg the separation of shadow and object implies that the object is not resting directly on the ground plane—see Kersten et al 1997), they clearly fall into the expendable-cue category as defined in the introduction. Therefore, it makes sense that artists tend to include fewer cast shadows in their work: if human observers do not easily identify the absence or inconsistencies of cast shadows, then why not exclude them when they become problematic?

3 Experiment 2

In experiment 2, we addressed the second half of our research hypothesis: that cast shadows in static, pictorial art are inherently ambiguous (Cavanagh and Leclerc 1989), and that movement will increase the ability of a human observer to distinguish an object's pigment from any shadows that have been cast onto the object's surface. This is because shadows will move differently than pigment.

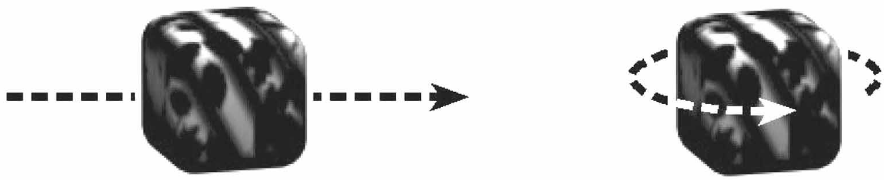
3.1 *Method*

3.1.1 *Participants.* The twenty participants from experiment 1 also participated in experiment 2. Two additional undergraduate participants (one male and one female, age 20 years) also took part in this experiment for a total of twenty-two participants.

3.1.2 *Materials.* We constructed 48 short Quicktime movies (gray-scale, 320×240 NTSC MPEG1 resolution, 90 frames at 30 frames s^{-1} for a total length of 3 s) using Maxon Cinema 4D R8. The movies were designed such that the viewer would have to discriminate between patterns that were on the surface of an object and patterns that were cast upon the surface by static shadows. The object we used was a cube with rounded corners. The participants first viewed the cube in a stationary position, then moving in a scene. Individual movies differed along four factors: pattern, penumbra, speed of movement, and type of movement. Two patterns were used: one was a simple striped pattern, the other was a leafy pattern (see figure 5). For each trial, the scene could show either a striped cube with a leafy shadow pattern cast upon it or a leafy cube with a striped shadow cast upon it. Half of the trials used the 'leafy' cube, and the other half used

Figure 5. The two different cubes used in experiment 2.

the 'striped' cube. Each of the patterns, whether pigment or shadow, could either be rendered with a crisp edge or with a penumbra that is often characteristic of cast shadows. There were three movement speeds: the cube either moved a substantial distance through the display (fast movement); the cubed moved a small distance very slowly (slow movement); or the cube was presented in two dichotomous positions without a smooth movement transition between them, inducing apparent motion (dichotomous movement). Lastly, the type of movement for half of the trials was a lateral translation, orthogonal to the viewing direction, while the other half of the animations consisted of rotations around the vertical object axis (see figure 6). The amounts of movement in the translation condition corresponded to 2 times the horizontal dimension of the object (slow movement) or 4 times the horizontal dimension (fast movement). In the rotation condition, the rotation was either 20° (slow condition) or 180° (fast condition). The amount of change in the dichotomous-motion condition was matched to the slow movement.



(a) (b)

Figure 6. Examples for the two main types of movements: (a) translation of the cube (always from left to right); (b) rotation of the cube around its vertical axis (always counterclockwise).

3.1.3 *Procedure.* Participants were tested individually in front of a computer monitor. The experimenter opened a movie to the first frame (still image) and asked the participant to indicate which of the two patterns corresponded to the pigment on the cube by naming the cube either as 'striped' or 'leafy'. The participants were encouraged to choose one of the two, but they could indicate that they were just guessing. After the participants had indicated their choice, the animation was started. After viewing the animation, the participants were again asked to indicate what pattern was on the cube. The order of presentation of the 48 trials was randomized for each participant.

3.2 *Results*

The results of this experiment are very straightforward. Just as expected, the guessing performance of participants given the initial, static view (mean = 46% correct, SD = 7.5%) was close to the expected value of 50% and, if anything, participants guessed more often incorrectly than correctly. This mainly illustrates that our stimulus set and experimental setup did not provide observers with clues about the identity of the cube. If designed correctly, the identification of the identity of the cube should be computationally impossible.

In the translation condition, the post-animation identification performance was 52% (SD = 19%) in the dichotomous condition, 86% (SD = 23%) in the slow movement

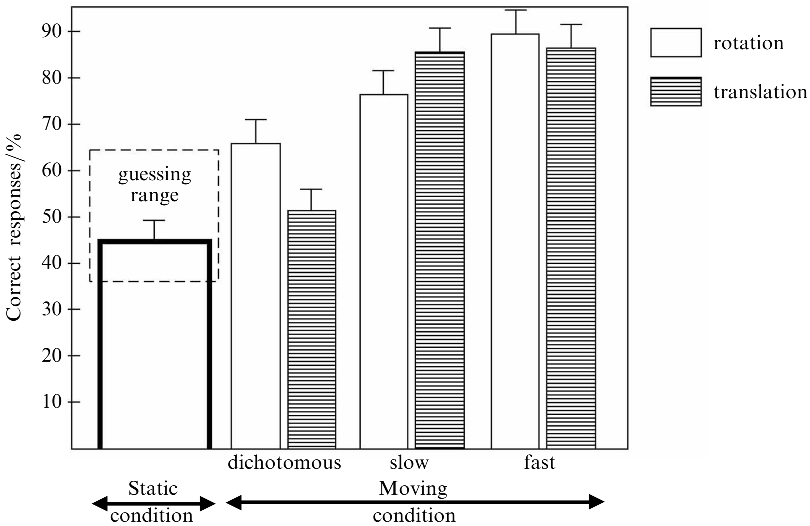


Figure 7. Mean correct identification rates for the pigment associated with the cube stimulus for the different conditions. The static view response corresponds to the initial perception before any animation of the stimulus. The guessing range for static images is based on the central 95% interval of the binomial distribution (0.5, 48). The dichotomous, slow, and fast conditions increase in their contiguity and the extent of the movement displayed in the animation. Error bars indicate standard errors.

condition, and 88% (SD = 22%) in the fast movement condition. For rotation trials, the performance increased from 67% (SD = 23%) in the dichotomous condition, to 78% (SD = 22%) in the slow movement condition, and 90% (SD = 21%) in the fast movement condition (see figure 7).

The results for each participant in the 6 different conditions were submitted to a two-factorial repeated-measures ANOVA with the factors type of movement (translation versus rotation) and movement continuity (dichotomous, slow, and fast). As in experiment 1, the degrees of freedom were Greenhouse–Geisser corrected for nonsphericity. The continuity and speed of the animation had a significant effect on identification rates ($F_{2,42} = 28.35$, $p < 0.001$), while the type of movement did not show a significant effect ($F_{1,21} = 1.013$, $p = 0.326$). The interaction between both factors was significant ($F_{1.9,39.2} = 5.49$, $p = 0.009$). The only condition in which the 95% confidence interval of the identification rate included the theoretical 50% guessing rate was for the dichotomous-translation animations. Paired-sample t-tests for the three different levels of movement continuity and speed show that slow, continuous movement was generally more efficient in producing correct identifications ($t_{21} = 4.98$, $p < 0.001$), and that faster and more extensive movements produced even larger benefits than the slow movement condition ($t_{21} = 3.03$, $p = 0.006$).

The manipulation of the penumbra for both the shadow and texture patterns produced different guessing rates in the static viewing condition. The cube's pigment was correctly identified on average 52% (SD = 15%) if the pigment did not have a penumbra, and only 39% (SD = 14%) if the pigment did show a penumbra. The presence of a penumbra in the shadow pattern led to a slightly better guessing rate of 48% (SD = 12%) than when the shadow pattern did not show a penumbra (mean = 43%, SD = 13%) (see figure 8). After viewing the animation, the guessing rates in all conditions were between 85%–86%, no matter which penumbra combination was presented.

Only the data for the static view condition were further analyzed since the homogenous means in the post-movement condition did not show any reliable differences.

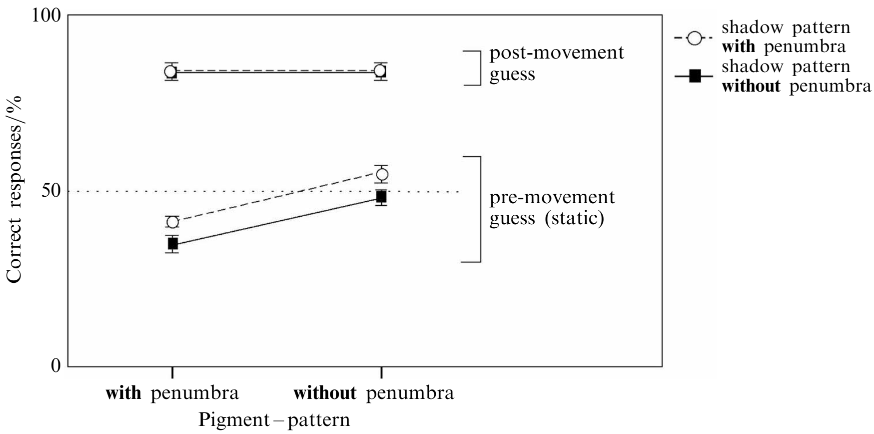


Figure 8. Mean correct identification rates for the pigment associated with the cube stimulus in relation to the presence or absence of a penumbra for the static view condition and after the presentation of continuous movement. Error bars indicate standard error.

The static-view data were submitted to a two-factorial repeated-measures ANOVA with the two factors pigment – penumbra and shadow – penumbra (both present versus absent). The results showed a significant effect of the penumbra attached to the pigment ($F_{1,31} = 6.04$, $p = 0.023$), while neither the presence of a penumbra of the shadow ($F_{1,21} = 1.21$, $p = 0.285$) nor the interaction between both factors ($F_{1,12} < 1$) was significant.

3.3 *Discussion*

The results of experiment 2 convincingly illustrate the importance of movement in disambiguating an object's pigment from shadows that fall onto the object. Even though two dichotomous views of an object can provide enough information to differentiate the two views in static depictions, human observers cannot necessarily integrate these views over time. This was most evident in the translation condition, which showed no improvement when viewed dichotomously, even though we had made sure that the shadow patterns were clearly different in the two views (eg a phase shift in the striped shadow pattern). There was a slight improvement for the rotation condition, which probably can be explained by the change in orientation of the patterns after the rotation. It might also be due to the spatial proximity of the two presentations in contrast to the spatial separation of the two views in the translation condition.

The addition of continuous motion in the slow movement and fast movement conditions allowed observers to reliably discriminate between pigment and shadow in the animations. Faster and more extensive movements led to the best discrimination ability of about 89% correct. All but one participant showed much improved discrimination performance in these two conditions. Even though this finding is not groundbreaking, it nicely illustrates the power of movement information that is not available in static media.

A second question that was addressed by experiment 2 was the role of penumbras for the discrimination of pigment and shadow. We factorially crossed the presence of a penumbra in either the pigment or the cast shadow, and the results are generally in the direction we expected. The presence of a penumbra biases observers to the assumption that this part of the image is a shadow, and not pigment attached to the object. In the two extreme cases, where either the pigment had a penumbra and the shadow did not, or where the shadow had a penumbra but not the pigment, we found a bias of approximately 16%. The important role of the penumbra was also stated by a large number of our participants when asked for their guessing strategy. However, the effect of the penumbra was smaller than we had anticipated. This might have been a factor of the factorial design and the experience of participants that the penumbra did not correlate with the pigment/shadow distinction in this study. On the basis of our results, we might thus underestimate the role of a penumbra in other, nonexperimental settings.

What seems clear, however, is that the role of the penumbra is trumped by movement. The guesses rendered by our participants in the two continuous movement conditions do not show any indication of being influenced by the presence or absence of a penumbra. The use of a penumbra to represent shadow thus seems limited to static media.

4 General discussion

How expendable are cast shadows for pictorial work? We approached this question from two different sides. The first experiment dealt with the necessity for inclusion: how critical are cast shadows to the interpretation of a scene, or how tolerant are observers to incongruities? The second experiment dealt with one particular motivation for the exclusion of cast shadows in static media, namely that cast shadows have the potential to act as a kind of camouflage, creating distracting ambiguities. The results of our two experiments indicate that inclusion of cast shadows may not be particularly critical to the legibility of a pictorial scene, and that exclusion may actually benefit the overall clarity of a work. Thus, we conclude that cast shadows may be considered expendable on both counts: not especially useful as well as potentially distracting.

Experiment 1 revealed that observers were relatively insensitive to local inconsistencies in cast shadows. Results demonstrated that in both conditions (incongruous orientation of a cast shadow and incongruous elevation of an object) the incongruities were not immediately apparent to the viewer. The clear trend towards longer search times with increasing number of objects indicates that viewers were searching serially for cast-shadow incongruities. This result fits nicely with the results of Ostrovsky et al (2001), who showed that inconsistent illumination of a scene was not easily discernible from inconsistent shading of objects. Similarly, the fact that viewers performed serial searches in the elevation condition was surprising, given the strong elevation cues revealed in the Kersten et al (1997) ball and shadow demonstration. However, it is worth observing, that unlike the ball and shadow demonstration, our stimuli were static.

Experiment 2 demonstrated that in cases where shadow and pigment are ambiguous, and no other cues are available, correct resolution from a static scene can be little more than guesswork. As soon as there is movement, the possibility for disambiguation appears. However, the success of this disambiguation is dependent on the nature of the movement. The longer and more continuous the movement, the more certainty and accuracy gained in discerning what is shadow and what is pigment. With less movement information, such as in the case of the slow-movement conditions, participants were less accurate and expressed less confidence in their choices. The least accurate responses were found in the dichotomous conditions, where there was a change in position or orientation of the target object, but no continuous motion of any kind. In the dichotomous conditions, participants performed only slightly better than chance.

After experiment 2, we queried participants about their strategies for discerning shadow pattern from pigment pattern in the initial static view. All indicated that they had been more or less guessing, but many had reasoned (at least in the beginning) that the shadows would be likely to be the patterns with penumbras. This bias was also reflected in our results, but only if no movement information was present. Once continuous motion was added, the small biasing effect of the penumbra vanished. One extremely attentive participant figured out that there were minute variations in the placement of the patterns on the cube depending on whether or not it was shadow or pigment.

She used these to successfully determine the correct response for the last 10 static views. We were aware of these minute differences, as they were largely due to the technical difficulty in perfectly matching up a pattern projected from a light source with a pattern on the surface of the cube. However, it seems clear from the data that participants could not use this information consistently.

It would seem then, from the results of these experiments, that artists have good reason to avoid the ambiguities that cast shadows can introduce into a piece if the penalty for excluding them is so negligible. This does not mean that cast shadows are always undesirable distractors, but rather that discretion needs to be exercised in their placement, so that they do not camouflage, flatten, or confuse the underlying surface. Da Vinci cautioned that care should be taken with shadows that fall across bodies, that the shadows should follow the contours of the form (MacCurdy 1938). Obviously, there are many strategies that can be used to distinguish shadow from pigment in static media. As seen from experiment 2, viewers tend to assume that a fuzzy edge is compatible with a shadow's penumbra rather than pigment. One participant put it this way: "You don't usually see those fuzzy edges on many things that aren't shadows". And just as in the real world, much of what is 'read' as shadow in a picture will depend on what is most likely based on the circumstances of the scene.

However, the aim of this paper was not to discuss how shadows should or should not be utilized in static media; artists have managed to work out this problem quite well without benefit from empirical studies by perceptual researchers. This paper began with the premise that the records left by artists through their work can offer clues about the way that human perceptual systems utilize various kinds of sensory information. One clue that seems to emerge from pictorial work is that the information conveyed by cast shadows is, at least in part, different from the information conveyed by attached shadows. Attached shadows have been fully adopted for use in pictorial production in the form of shading, to lend the illusion of form and three-dimensionality to two-dimensional surfaces. Cast shadows, on the other hand, have been used much more sparingly in pictorial art, and, as we have seen from the results of these two small studies, they may be considered fairly expendable for the legibility of a static scene.

This study demonstrates that continuous motion can help disambiguate cast shadow from pigment, but that dichotomous motion or apparent motion is of little help. Similarly, in the change-blindness studies by the second author, viewers were insensitive to gross incongruities of cast shadow in two differing views, which is an essentially dichotomous situation. It would appear then, that cast shadows, unlike attached shadows, are primarily conveyors of transient or temporal information, and that the utility of cast shadows is considerably diminished in any kind of static situation.

Acknowledgments. Both authors contributed equally to this work. We decided on the order of authorship mainly on the first author's initial motivation and continuing drive to conduct this study. We wish to thank Keith Moore, Derek Schorsch, Thomas Sneed, and Mark Yama for their assistance in running the experiments and Greg Turner-Rahman for his advice on 3-D animation programs. We also wish to thank the two anonymous reviewers of previous versions of this manuscript for their thoughtful comments.

References

Behrens R R, 1998 "Max Wertheimer and Pablo Picasso: Gestalt theory, Cubism and camouflage" *Gestalt Theory* **20** 111 – 118

Casati R, 2003 *The Shadow Club* (New York: Knopf)

Cavanagh P, Leclerc Y C, 1989 "Shape from shadows" *Journal of Experimental Psychology: Human Perception and Performance* **15** 3 – 27

Cott H B, 1996 "Animal form in relation to appearance", in *Aspects of Form* Ed. L Law Whyte (Bloomington, IN: Indiana University Press) pp 121 – 156

Enns J T, Rensink R A, 1990 "Influence of scene-based properties on visual search" *Science* **247** 721 – 723

Gombrich E H, 1995 *Shadows: The Depiction of Cast Shadows in Western Art* (London: National Gallery Publications)

Janson H W, Janson D J, 1984 *Janson's Story of Painting: From Cave Painting to Modern Times* (New York: Harrison House/Harry N Abrams)

Kersten D, Mamassian P, Knill D C, 1997 "Moving cast shadows induce apparent motion in depth" *Perception* **26** 171 – 192

Lewis-Williams D, 2002 *Mind in the Cave: Consciousness and the Origins of Art* (London: Thames & Hudson)

Luckiesh M, 1965 *Visual Illusions: Their Causes, Characteristics and Applications* (New York: Dover)

MacCurdy E, 1938 *The Notebooks of Leonardo da Vinci* volume II (New York: Reynal & Hitchcock)

Mamassian P, Knill D C, Kersten D, 1998 "The perception of cast shadows" *Trends in Cognitive Sciences* **2** 228 – 295

Mossy Oak, 2003 *Research and Development* (retrieved July 25, 2003, from http://www.mossyoak.com/camoguide/r&d.asp)

Ostrovsky Y, Cavanagh P, Sinha P, 2001 "Perceiving illumination inconsistencies in scenes" AI Memo 2001-029, November 2001, Massachusetts Institute of Technology, Cambridge, MA, USA

Ramachandran V S, Hirstein W, 1999 "The science of art: A neurological theory of aesthetic experience" *Journal of Consciousness Studies* **6** 15 – 51

Soby J T, 1969 *The Early Chirico* (New York: Arnos Press)

Thayer A H, 1909 *Concealing-Coloration in the Animal Kingdom* (New York: Macmillan)

Werner S, Thies B, 2000 "Is 'change blindness' attenuated by domain-specific expertise? An expert – novices comparison of change detection in football images" *Visual Cognition* **7** 163 – 173

Zeki S, 2001 "Artistic creativity and the brain" *Science* **293** 51 – 52

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 33 2004

www.perceptionweb.com