

---

# The hollow-face illusion: Object-specific knowledge, general assumptions or properties of the stimulus?

---

Harold Hill

Human Information Science Laboratories, ATRi, Keihanna Science City, 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619 0288, Japan; e-mail: hill@atr.jp

Alan Johnston

Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK

Received 23 September 2005, in revised form 22 March 2006; published online 31 January 2007

---

**Abstract.** The hollow-face illusion, in which a mask appears as a convex face, is a powerful example of binocular depth inversion occurring with a real object under a wide range of viewing conditions. Explanations of the illusion are reviewed and six experiments reported. In experiment 1 the detrimental effect of figural inversion, evidence for the importance of familiarity, was found for other oriented objects. The inversion effect held for masks lit from the side (experiment 2). The illusion was stronger for a mask rotated by 90° lit from its forehead than from its chin, suggesting that familiar patterns of shading enhance the illusion (experiment 2). There were no effects of light source visibility or any left/right asymmetry (experiment 3). In experiments 4–6 we used a ‘virtual’ hollow face, with illusion strength quantified by the proportion of noise texture needed to eliminate the illusion. Adding characteristic surface colour enhanced the illusion, consistent with the familiar face pigmentation outweighing additional bottom–up cues (experiment 4). There was no difference between perspective and orthographic projection. Photographic negation reduced, but did not eliminate, the illusion, suggesting shading is important but not essential (experiment 5). Absolute depth was not critical, although a shallower mask was given less extreme convexity ratings (experiment 6). We argue that the illusion arises owing to a convexity preference when the raw data have ambiguous interpretations. However, using a familiar object with typical orientation, shading, and pigmentation greatly enhances the effect.

## 1 Introduction

The hollow-face illusion is a well-known example of depth inversion occurring with a real object under a wide variety of viewing conditions (Gregory 1970). The key, unanswered, question remains—why? Despite the availability of unambiguous cues to relative depth, for example binocular disparities and motion parallax, coupled with explicit knowledge of the true shape of the mask, we have the compelling and obligatory impression of a convex face with the nose sticking out. Contrary to previous assertions (Georgeson 1979; Hill and Bruce 1993), this is not a case of cue conflict. With a real mask all of the usual cues are present, determined by and consistent with the actual, concave, surface. Nor is the illusion a case of bistability, with alternative perceptual interpretations alternating—beyond a certain distance the convex percept dominates and no effort of will can cause it to flip. A satisfactory answer to the question ‘why?’ would have implications well beyond the illusion itself. In particular, the illusion will provide pointers relevant to the classic problem in vision: how do we see a three-dimensional world from a two-dimensional retinal image?<sup>(1)</sup>

<sup>(1)</sup>The reduction in dimensionality from the distal to the proximal stimulus is only one of the ways of expressing the ‘inverse problem’. Given the difficulty in even conceptualising a 3-D retina, it can be argued that the real problem is the limitations of a single line-of-sight, at least for a stationary, one-eyed observer (Rogers 2003, pp 29, 572–573; Howard and Rogers 1995). This gives rise to the fundamental ambiguity of vision, whereby an infinite number of scenes can give rise to a particular retinal image (Kersten et al 2004). However it is expressed, attempts at constructing artificial vision systems have served to emphasise that there is a problem to be solved. What is remarkable about the human visual system is how rarely it breaks down. We concur with Gilchrist (2003) in believing that such occasional failures are the bugs that can help to reveal underlying algorithms and assumptions.

Existing explanations of the hollow-face illusion reflect theoretical standpoints on this broader question. We characterise these explanations as lying on a continuum. At one end are top-down explanations that emphasise prior object-specific knowledge about the familiar convex shape of faces. In the centre are mid-level accounts invoking general knowledge about the world embodied in the form of assumptions applied in the interpretation of the raw sensory data. Finally, at the far end, are purely bottom-up accounts framed solely in terms of the sensory data.

For many, the hollow-face illusion is one of the strongest pieces of evidence for the importance of top-down knowledge in perception (Georgeson 1979; Gregory 1970, 1997; Heard and Chugg 2003; Klopfer 1991; Leweke et al 1999, 2000; Papatomas and Bono 2004; Schneider et al 2002; Wolf and Wolf 1990). Schroeder (1852), for example, states that the illusion only works if the hollow shape resembles a familiar 3-D object.

In this view “top-down knowledge of faces is pitted against bottom-up signaled information” (Gregory 1997, page 1122) and knowledge wins. The illusion is classified as belonging to the class of ambiguous illusions and its cause ascribed to knowledge—that is to cognitive not physical factors. Within this theoretical framework, perception can be seen as analogous to the scientific process of testing hypotheses against the available data (Gregory 1980). In the case of the hollow face there are (at least) two competing hypotheses—that a particular object is convex or concave. Gregory argues that it is the bias for seeing faces as convex (Gregory 1997) that leads to the compelling impression of the familiar percept, a convex face with the nose pointing towards you. The top-down knowledge involved in these explanations is object-specific and associated with differences in ‘bias’, ‘probability’, or ‘familiarity’ associated with convex, as opposed to concave, faces. In this case, the very existence of the illusion would be evidence that our representations of faces are inherently three-dimensional. Arguments citing object knowledge have also been invoked to explain the failure of depth inversion under pseudoscopic viewing conditions, that is with the left and right binocular images exchanged (Deutch et al 1990; Gregory 1970; van der Enden and Spekrijse 1989; Wheatstone 1838, 1852). However, the pseudoscopic case differs in that the monocular cues remain consistent with the convex percept while the binocular disparities alone imply concavity (although some monocular cues, in particular shading, are consistent with both convex and concave percepts as discussed in the paragraphs on bottom-up explanations below). Lastly, reduced susceptibility to the illusion reported to be associated with certain drugs and mental disorders has been attributed to impairments in processing of this top-down knowledge (Emrich 1989; Leweke et al 1999, 2000; Schneider et al 2002).

The primary evidence for the importance of object-specific knowledge or familiarity is the effect of figural inversion—rotation by  $180^\circ$  in the image plane. For example, a hollow mask rotating about its vertical axis reverses to appear as a face rotating in the opposite direction more often when it is presented upright than when it is upside-down (Klopfer 1991). This is attributed to the ‘differential familiarity’ of upright and inverted views and interpreted as evidence that familiarity can affect the perception of rigid motion. The effect of inversion cannot be explained in terms of stimulus properties, as these remain the same regardless of orientation. In Klopfer’s work no effect of orientation was found with static photographs, leading the author to argue that the cognitive effect is on motion rather than on the perception of three-dimensional shape. However, evidence has been found in subsequent work that figural inversion directly affects the perception of the static mask as convex or concave (Hill and Bruce 1993, 1994; Papatomas and Bono 2004). Specifically, the illusion persists at shorter distances and for a greater proportion of the time when the mask is presented upright rather than upside-down. Again, this cannot be explained in terms of stimulus properties and is not true for objects without a characteristic orientation (Hill and Bruce 1994).

The effect of orientation appears more pronounced for faces than for scenes shown in reverse perspective (Papathomas and Bono 2004). In the experiments on rigid motion, an effect of inversion was only found for the face and not for the ‘jar’ or either of the unfamiliar objects used (Klopper 1991). In experiment 1 we test the effect of figural inversion for a variety of objects.

Evidence against the importance of object knowledge comes from reports of depth inversion occurring for objects where convex and concave alternatives are not differentially familiar, including the Mach card (Mach 1886/1959), complex nonsense objects, or crumpled newspaper (Deutch et al 1990), and the ‘hollow-potato’ (Hill and Bruce 1994; Johnston et al 1992). These examples have been taken as evidence for a general ‘convexity preference’ or assumption in the perception of ambiguous stimuli, including the hollow-face illusion (Ramachandran 1995). This can be construed as the general knowledge that objects are lumps of matter and therefore globally convex.

A convexity preference is characteristic of the kind of general (rather than object-specific) knowledge associated with our second class of explanations. Another example is Brewster’s explanation that “we judge the forms of bodies by the knowledge we have acquired of light and shadow” (Brewster 1826, 1832; Wade 1983). Most of the examples given and cited by Brewster involve the use of arrangements of lenses that ‘invert’ the image of the object, that is rotate it 180°, causing its pattern of shading to be left–right reversed with respect to the true position of the light. In this respect, Brewster’s explanation follows Rittenhouse (1786). However, Brewster adds examples where such lenses are not necessary, including the ability “to raise a complete hollow mask of the human face into a projecting head” (Wade 1983). One assumption that is often invoked in the interpretation of shading and shadow is an assumption that light comes from above (Bermbaum et al 1983; Ramachandran 1988). Thus, the same shaded figure can reverse between convex and concave when it is figurally inverted. In this case, the apparent light source appears to remain in the same place, above with respect to the viewer, while the pattern of shading reverses. In these cases, the preference for light-from-above appears to be dominating any preference for convexity. However, in the case of the hollow-face illusion, this does not necessarily happen, and a picture of the hollow face can still look convex when it is inverted and the light source is perceived as being from below (Gregory 1973). We test the effects of the direction and visibility of the light source in experiments 2 and 3 and the shading-related effect of photographic negation in experiment 5.

Wheatstone criticises Brewster’s explanation in terms of knowledge of light and shadow. Instead he emphasises that for cameos and intaglios of the same object (which are convex and concave, respectively) “the projection of either on the retina is sensibly the same” (Wheatstone 1838). While he admits the influence of “accessory circumstances” and “previous knowledge” of shadows he argues that the “real cause” of the illusion arises from “our more perfect means of judging being absent” (Wheatstone 1838). This introduces a third class of explanation, those that explain depth inversion in terms of bottom–up stimulus properties alone. Shading in particular is inherently ambiguous, or at least its interpretation is (Gregory 1973): there are always two possible interpretations, convex or concave with opposite apparent lighting directions (Horn 1977). One of the first reports of depth reversal suggests “omnes has fallacies per umbram oriri” (Gmelin 1744)—“all those fallacies were owing to shade” (Brewster’s 1826 translation cited on page 58 of Wade 1983).

We might expect some monocular cues to be unambiguous and to provide Wheatstone’s “perfect means of judging”. For example, monocular-perspective texture gradients, as opposed to object knowledge, have been suggested as the reason why pseudoscopically presented faces do not depth-invert (van der Enden and Spekrijse 1989). However, in the case of the hollow face, this cue would veridically indicate concavity and so cannot

be the explanation (Deutch et al 1990). In addition, both texture gradients and perspective are themselves open to ambiguous interpretations. For example, a wire-frame cube appears to be a truncated pyramid when seen as depth-inverted (Shopland and Gregory 1964). Both the cube and the pyramid will project the same retinal image under certain viewing conditions. ‘Reverspective’ paintings deliberately exploit the ambiguity of both texture and perspective, making use of carefully constructed perspectives that project images contradictory to their actual depth—near surfaces being made proportionately smaller and far surfaces larger (Papathomas 2002; Papathomas and Bono 2004; Wade and Hughes 1999). Gradients of anisotropic texture elements (that is elements that get bigger with actual viewing distance, reversing the gradient of diminishing projected size normally associated with isotropic, equally sized, texture elements viewed in perspective) and ‘shadows’, both defined by pigment and consistent with the reversed depth, are added to strengthen the effect. In these paintings many mid-level rules/assumptions, including the isotropy of surface texture elements and rectilinearity, are deliberately violated. However, in the case of the hollow face, these cues and their associated rules are not artificially manipulated and might be expected to provide disambiguating cues. We investigate the effects of these variables in experiment 4 by comparing orthographic and perspective projection with and without surface-colour information.

Like others, notably Helmholtz and Mach (Yellott 1981), Wheatstone (1838) denies the possibility of binocular inversion, asserting that “when the cameo or intaglio is seen with both eyes, it is impossible to mistake an elevation for a depression”. He also attributes later reports of such observations to suppression of one of the images (Wheatstone 1852). Schroeder (1852), on the other hand, reports that the minimum distance for the illusion is larger when viewed with both eyes, a claim consistent with more recent experimental evidence (Hill and Bruce 1993). Underlying this difficulty in accepting the possibility of binocular depth inversion is the fact that, while many properties of the monocular image are potentially or inherently ambiguous, binocular-disparity information is unequivocally unambiguous with respect to concave and convex interpretations.

How then can we account for the hollow-face illusion, a clear example of *binocular* depth inversion? This remains a fundamental problem for bottom-up explanations. It is not, contrary to Wheatstone’s suggestion, a case of monocular suppression—when viewed from directly in front, the nose points straight towards us (towards our ‘cyclopean eye’) and not to the left or the right eye. Neither is it a case of mental reversal of left and right images, ‘mental pseudoscopy’ (Yellott and Kaiwi 1979): a simultaneously presented stereogram is seen in the correct depth at the same time as the illusory depth-reversed face is perceived (Yellott and Kaiwi 1979). Similarly, the illusory rotation of Pulfrich’s pendulum is in the direction expected (as defined by the light-attenuated eye) at the same time as an attached mask is perceived as convex (Yellott 1981). In experiment 6 we investigate the effect of binocular disparities through scaling the depth of the mask.

Another potentially unambiguous cue, closely related to stereoscopic disparities, is motion parallax. If the observer moves parallel to the plane of the face, the differences in motion on the retina(s) between the edges of the face and the nose provide potentially unambiguous depth information that the mask is hollow. However, instead of the illusion reversing in depth and being perceived veridically as hollow, the illusory convex face is perceived to rotate in the direction of observer movement at twice the rate of angular rotation. This apparent movement can be explained geometrically. The visible proportion of each half of a mask or face changes with viewing angle as a function of foreshortening. For a mask, the proportion of the half of the face nearest the viewer decreases, whereas the reverse is true of a convex face. Changes in viewing angle can result from either rotation of the mask or movement of the viewer.

---

Thus, when a mask is viewed from an angle, the change in proportions resulting from the change in angle relative to the mask is consistent with a convex face rotated the same angle relative to the observer's new line of sight. The total apparent rotation of the illusory mask is therefore the sum of the two rotations, twice the viewing angle. The low-level motion on the retina is the same for the stationary concave mask and a doubly rotating convex face, and the motion, rather than serving to disambiguate the illusion, is reinterpreted in a manner consistent with the percept of a convex face. This is a beautiful example of 'perceptual explaining away' (Kersten et al 2004).

In summary, Wheatstone is probably correct to argue that there is no "perfect means of judgment" available, at least in the case of monocular viewing. However, naturally occurring texture gradients would normally be expected to provide reliable cues, and binocular disparities are not ambiguous. Any bottom-up account needs to explain what happens to the binocular disparity information and also why the 'wrong' solution wins in the case of ambiguous cues.

Two recent papers have both reported work designed to test if the 'action system' is susceptible to the hollow-face illusion (Hartung et al 2005; Kroliczak et al 2006). Clearly, if the action system were not susceptible to the illusion, this would imply that there is sufficient disambiguating information available. We would still need to explain why our awareness is of a convex face. Currently, the evidence is contradictory, with one paper indicating that the action system is not susceptible to the illusion (Kroliczak et al, in press) while the other argues that it is (Hartung et al 2005). This issue, while fascinating in itself, is not directly relevant to this paper where the focus is limited to perception and the experiments did not involve action. Of relevance here is that one of these papers made use of a real object (Kroliczak et al, in press) while the other used virtual stereoscopic presentation (Hartung et al 2005)—another possible contribution to the apparently contradictory evidence. We used real objects for experiments 1–3, but a virtual stereoscopic display for experiments 4–6, discussing possible differences in the introduction to experiment 4.

Both action-based papers report a 'compression effect', with the illusory face appearing flatter than an equivalent actual face or veridically perceived hollow mask. This effect does not appear to have been noted in the literature on illusions of depth reversal before. The effect is interpreted as reflecting the combination of disparity information with familiarity and other cues to depth (Hartung et al 2005), or the operation of Emmert's law (Kroliczak et al, in press). If it results from misperceived distance of the illusion, the flattening may also reflect the fact that size scales with  $1/d$  whereas disparity scales as  $1/d^2$ , a fact well-known to cause illusory flattening (for review, see Howard and Rogers 2002). The effect of the relative depth of the illusion is considered in experiment 6.

In this paper we attempt to provide experimental evidence relevant to a number of the issues raised above concerning the cause of the perceptual illusion. This evidence is intended to facilitate a choice between the alternative classes of explanation outlined: object-specific familiarity, general knowledge embodied as assumptions, and purely bottom-up explanations. The experiments were intended to elucidate the role of familiarity in perception, particularly in the integration of cues to depth. Experiment 1 uses figural inversions with objects other than the face to investigate the role of differential familiarity in generating the illusion. In experiment 2, we also investigate familiarity, seeking to separate the effect of figural inversion from the effect of familiar lighting direction. In experiment 3, we test whether the visibility or otherwise of the light source is critical to the illusion. Experiments 4–6 use a virtual hollow-face illusion to give additional control over stimulus properties. A noise texture [for an example of the texture, see figure 8 or the additional materials online (<http://www.perceptionweb.com/misc/p5523/>)] is added to the virtual mask in order to control the stereoscopic information

available. Confidence ratings related to convexity are then recorded as a function of the proportion of noise texture added under the different experimental conditions. In experiment 4, the effect of adding natural surface pigmentation is investigated, as is the effect of orthographic as opposed to perspective projection. In experiment 5, we make use of photographic negation of the images to disrupt shading information, while in experiment 6 the effect of the absolute depth of the mask is investigated. Finally the implications of the results for explanations of the illusion are considered in the general discussion.

## 2 Experiment 1

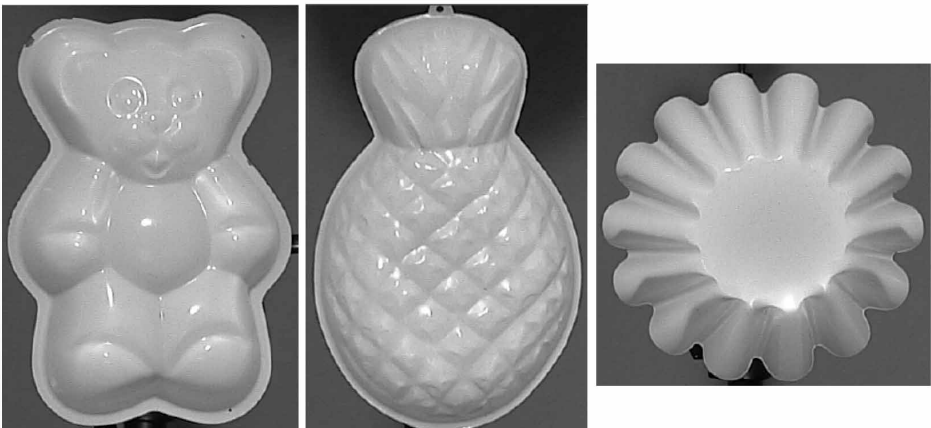
This experiment uses the effect of figural inversion to investigate the role of familiarity in illusions of depth reversal. As discussed in the introduction, previous evidence suggests that the more familiar upright view of a face produces a stronger illusion than the relatively unfamiliar, inverted face (Hill and Bruce 1993, 1994; Klopfer 1991; Papathomas and Bono 2004). In this experiment we wished to test whether there are effects of inversion for other objects.

Previous work has not shown significant effects of inversion for other objects, for example a jar (Klopfer 1991) or a scene (Papathomas and Bono 2004). There is also no effect of inversion for previously unfamiliar objects without a familiar orientation (Hill and Bruce 1994; Klopfer 1991). This suggests that the effect of familiarity might be limited to, or at least strongest for, faces. Further, an inverted face produces an illusion of similar strength to an unfamiliar object, suggesting that the effect of familiarity may be all-or-nothing (Hill and Bruce 1994).

If the effects of familiarity were limited to faces, we would expect all other moulds, regardless of their orientation, to produce illusions of the same strength as the hollow-potato or inverted face. Alternatively, if the effects of familiarity are more general, objects with a 'differentially familiar' (Klopfer 1991) orientation should generate more powerful illusions when presented upright than inverted. Testing a number of objects in a number of orientations should also show whether the effects of familiarity are all-or-nothing or graded, according to whether the strengths in the different conditions appear to take two distinct values or a range.

### 2.1 Methods

**2.1.1 Materials.** The stimuli were three commercially available moulds with similar dimensions and made of metal but painted matte-white to reduce specularities. They were a teddy bear and a pineapple, both with a clearly defined vertical orientation, and a regular jelly mould with no clearly defined upright orientation (figure 1).



**Figure 1.** The three hollow moulds (teddy bear, pineapple, and regular jelly) used in experiment 1 as they appeared in the upright condition.

As can be seen from figure 1, specularities were not eliminated completely and would be expected to reduce the strength of illusions of depth reversal as binocular specularities can disambiguate convex from concave surfaces (Blake and Bülthoff 1990). The dimensions of the teddy, pineapple, and jelly moulds were 19 cm × 13 cm × 4 cm, 28 cm × 17 cm × 5 cm, and 18 cm × 18 cm × 6 cm for the height, width, and depth, respectively. These dimensions are similar to those of life-size facemasks (see experiments 2 and 3). The moulds were presented clamped to a retort stand and lit from below, a lighting direction known to enhance the strength of the hollow-face illusion lit from in front (Hill and Bruce 1993). Upright was defined arbitrarily for the regular jelly mould.

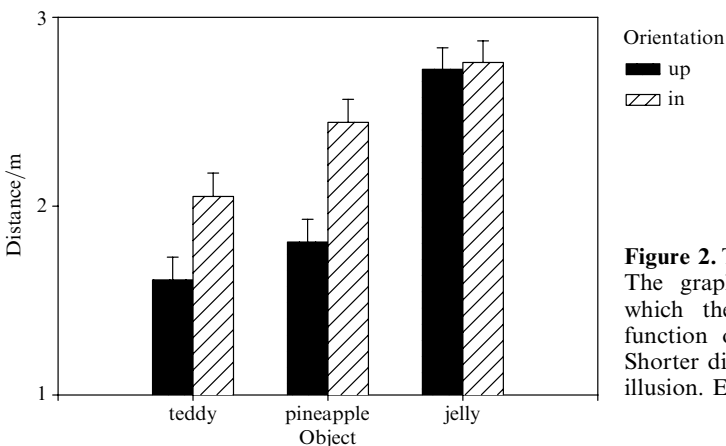
**2.1.2 Observers.** Twelve volunteers from the University College London Department of Psychology, all with normal or corrected-to-normal vision, took part in the experiment.

**2.1.3 Procedure.** Observers were introduced to the experimental situation and their task explained. They were asked to approach the hollow object from 4 m away and stop when their perception of it changed from convex to concave. They were encouraged to move from side to side to test their percept, as only the illusory percept appears to follow the observer. The three objects were then presented in each of their two orientations with order randomised. As a test of reliability, two measures were taken for each condition, both with observers approaching the moulds. The distance away at which their perception had changed was measured for each trial and the observer returned to the far end of the room before the next trial commenced. All viewing was binocular.

**2.1.4 Design.** The design was a 3 (object) × 2 (orientation) within-subjects factorial design. The dependent variable was the distance away at which observers reported that their perception had flipped from convex to concave. Two measures were taken for each condition, and measure number, whether first or second, was included as an additional factor in an initial analysis.

## 2.2 Results and discussion

An initial analysis showed no effect or interactions involving measure (all  $p$ s > 0.1) and so data were collapsed across this factor. A 3 (object) × 2 (orientation) ANOVA gave an object × orientation interaction ( $F_{2,22} = 4.7$ ,  $p < 0.05$ ). This interaction is presented graphically in figure 2. Analysis of simple main effects showed effects of orientation for the teddy bear ( $F_{1,24} = 7.6$ ,  $p < 0.05$ ), and the pineapple ( $F_{1,24} = 15.6$ ,  $p \ll 0.05$ ), but not for the jelly ( $p > 0.1$ ). There was also a simple main effect of object for both upright ( $F_{2,44} = 21.3$ ,  $p \ll 0.05$ ) and inverted ( $F_{2,44} = 7.6$ ,  $p \ll 0.05$ ) orientations.



**Figure 2.** The results of experiment 1. The graph shows the distance at which the illusion changed as a function of object and orientation. Shorter distances indicate a stronger illusion. Error bars show  $\pm 1$  SEM.

Pairwise comparisons ( $\alpha/6$ ) showed that for the upright orientation both the teddy ( $t_{11} = 7.8, p \ll 0.008$ ) and the pineapple ( $t_{11} = 4.5, p \ll 0.008$ ) produced stronger illusions than the regular jelly, but did not differ from each other ( $p > 0.1$ ). In the inverted orientation, the teddy produced a stronger illusion than the regular jelly ( $t_{11} = 3.8, p \ll 0.05$ ) and the pineapple ( $t_{11} = 2.3, p < 0.05$ ). The pineapple did not differ from the jelly ( $p > 0.1$ ).

The results replicate and extend previous findings in showing that the hollow-face illusion is not limited to faces. They also show that different objects generate illusions of different strengths, and that objects with a clearly defined orientation generate a more effective illusion in their familiar orientation. This is consistent with the general importance of familiarity as a factor affecting the perception of low-level depth cues. It appears from inspection of the graph that the effect of familiarity is not all-or-nothing as various levels are apparent. Statistically, the inverted teddy bear generates a weaker illusion than its upright equivalent but a stronger illusion than the regular jelly or inverted pineapple, showing that at least three distinct levels are possible.

The regular jelly mould produced the weakest illusion. This may have been because jelly moulds are not familiar objects, because they are familiar as *concave* objects, or because this mould was deeper than the others. The effect of inversion shows that depth alone cannot account for the pattern of results. Depth is explicitly examined in experiment 6. As expected, there was no effect of orientation on the regular jelly.

Although a face model made of the same material as the moulds used here was not available and previous testing conditions will inevitably have been different, informally comparing distances to previously reported data suggests that a hollow face produces an illusion at least as strong as any of the objects tested here—mean critical distance 1.25 m and 1.45 m (Hill and Bruce 1993, 1994). Interestingly, a reverse-perspective scene has been reported as being more compelling than a face, although it should be noted that the physical stimuli and cues available differed in a number of ways including size and the presence of painted shadows (Papathomas and Bono 2004). As with other face-perception tasks, the effect of inversion appears disproportionately large for faces compared to other objects (Valentine 1988)—mean effect for faces 1.17 m and 1.09 m (Hill and Bruce 1993, 1994) compared to 0.44 m for the teddy and 0.63 m for the pineapple. As noted above, a reverse-perspective scene (Papathomas and Bono 2004) and a jar (Klopper 1991) have not been found to show significant effects of inversion when tested previously.

Experiment 1 showed that other familiar objects besides the face also generate strong illusions of depth reversal. For objects with a characteristic orientation, inversion reduced the strength of the illusion, although not as much as for faces. This suggests that the effect of differential familiarity is not unique to faces but a general factor affecting the perception of low-level cues. The effect of familiarity does not appear to be ‘all-or-nothing’ as different versions of the illusion generate a range of values.

### 3 Experiment 2

In this experiment we sought to test whether there is a familiar lighting direction in the same way as there is a familiar orientation. The upright hollow face is more likely to appear convex than the inverted mask, independent of lighting direction (Gregory 1973; Hill and Bruce 1993; Wolf and Wolf 1990). This suggests that the effect of orientation is more important than that of lighting. When the mask is upright or inverted, the illusion seems to be stronger when light appears to be from above (Hill and Bruce 1993; Wolf and Wolf 1990), consistent with a light-from-above assumption. However, a weaker illusion has been reported for inverted faces lit from the chin (Sakurai et al 1985), perhaps because the pattern of shading for a face lit from above its forehead is more familiar than for a face lit from below its chin, as can be seen from comparing the 90° and 270° views in figure 3. The experiment was designed in part to test this





**Figure 3.** The hollow-face illusion used in experiment 2 shown in the four orientations and with the lighting used.

possibility. It also provided a test whether there would be an effect of inversion even when both orientations were neutral with respect to the light-from-above assumption.

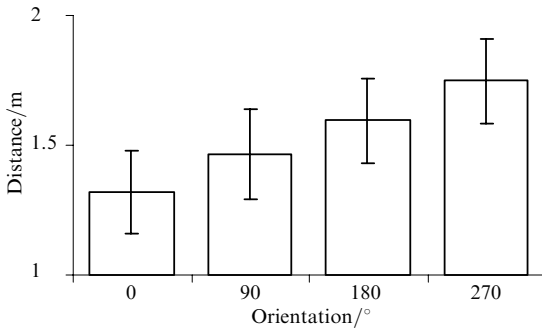
The hollow mask was shown in four orientations with lighting always from the right with respect to the observer (lighting was through the translucent mask, and so apparent and actual lighting directions coincide). If, as Brewster might argue (Brewster 1826, 1832), knowledge of lighting alone determines the strength of the illusion, no difference between conditions would be expected as the lighting remained unchanged. Only if the differential familiarity of upright and inverted faces is important, would we expect a difference between  $0^\circ$  and  $180^\circ$  orientations, and only if lighting direction relative to the head is important, would we expect a difference between  $90^\circ$  and  $270^\circ$  orientations.

### 3.1 Methods

The same method was used as for experiment 1 except that, in this experiment, one approaching and one retreating measure were used. The mask was made of white translucent plastic and its maximum dimensions were approximately  $21\text{ cm} \times 17\text{ cm} \times 7\text{ cm}$  (height  $\times$  width  $\times$  depth). The orientation of the mask was the independent variable, and both approaching and retreating measures were taken. Twelve observers recruited from ATR took part in the experiment as volunteers. We used a 4 orientation ( $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$ )  $\times$  2 direction (approaching, retreating) within-subjects design.

### 3.2 Results and discussion

As illustrated in figure 4 there was a main effect of orientation ( $F_{3,33} = 12.1$ ,  $p \ll 0.05$ ), but no effect of direction or any interaction. Planned  $t$ -tests showed significant differences between  $0^\circ$  and  $180^\circ$  ( $t_{11} = 3.6$ ,  $p \ll 0.05$ ), and  $90^\circ$  and  $270^\circ$  ( $t_{11} = 3.9$ ,  $p \ll 0.05$ ).



**Figure 4.** The results of experiment 2 showing the strength of the illusion as a function of orientation ( $0^\circ$  corresponds to upright—see figure 3). Short distances indicate a strong illusion and error bars show  $\pm 1$  SEM.

The first comparison shows that there is an effect of orientation, even when lighting is from the side and therefore neutral with respect to ‘above’ and ‘below’. The second comparison shows that lighting from above *relative to the head* produces a stronger illusion than lighting from below the head, even though the lighting remained neutral with respect to a ‘light-from-above’ assumption defined in terms of the observer’s own head (Howard 1990). The trend for the effect of direction was as expected, that is the initial percept was maintained causing a hysteresis-like overlap as a function of direction of movement (Hill and Bruce 1993)—approaching 151 (15) cm, retreating 155 (17) cm—but did not reach significance.

The results suggest that the object knowledge contributing to the illusion reflects both a familiar orientation and a familiar pattern of shading. Thus, the stored knowledge does not constitute a full object-centred abstract 3-D model, but an encoding that reflects prior experience with regards to both orientation and lighting. The object-centred pattern of shading visible in the  $90^\circ$  view of figure 3 is closer to this representation than the pattern visible in the  $270^\circ$  view, just as the  $0^\circ$  is more familiar than the  $180^\circ$  view. This suggests object-centred encoding of lighting distinct from the scene-centred effects reported previously (eg Tarr et al 1998). Lighting-dependent representations would also explain why a familiar, but not an unfamiliar, object triggers an appropriate interpretation of a two-tone image in terms of lighting and shadows (Moore and Cavanagh 1998). Just as view-invariant recognition may be supported by a limited number of viewpoint-specific representations (Perrett and Oram 1998), so a limited number of lighting-specific representations may be sufficient to support recognition from novel lighting directions (Belhumeur et al 1996; Shashua 1997).

### 4 Experiment 3

This experiment was designed to test whether the visibility of the light source affected perception of the illusion. In many of the early studies of the perception of cameos and intaglios reviewed in the introduction, perceived light source direction was believed to be critical (Wade 1983). If this is true, one might expect the illusion to be less strong when the light source is visible and inconsistent with the illusory percept. Two lighting directions were used, one from the left and one from the right, and so this experiment also provided a test whether these directions were inherently different. The mask used was a plaster mould illuminated from slightly in front and to the side so that, unlike the plastic mask illuminated from behind used in experiment 2, illusory and actual lighting directions are opposite each other (ie when lit from the left the upright mask appears lit from the right and vice versa).

#### 4.1 Method

The method was the same as in previous experiments and fifteen observers recruited from local Japanese universities participated. The face model used was based on an average 3-D head model constructed from 200 males and females (Vetter and Troje 1997)

which had been printed with a 3-D printer and housed with a black surround. The dimensions of the mask were 19 cm × 13 cm × 8 cm (height × width × depth). Observers were introduced to the illusion with both lights on and their task explained. Lights were hidden for 'invisible' trials using small screens. There was also a low level of ambient lighting in the room designed to minimise the visibility of shadow cues.

#### 4.2 Results and discussion

A 2 visibility (visible, invisible) × 2 side (left, right) × 2 direction (approaching, retreating) ANOVA showed a main effect of direction ( $F_{1,14} = 19.5, p \ll 0.05$ ) but no other main effects or interactions (all  $ps > 0.1$ ). Means (SEMs) were 176 (13) cm for approaching and 236 (16) cm for retreating trials. Trials with a visible and invisible light source had means of 204 (13) cm and 207 (14) cm, respectively. Values for left and right trials were 206 (13) cm and 207 (14) cm, respectively.

The lack of effects of the side or visibility of the lighting are null results and may reflect a lack of sensitivity of the method or, in the case of visibility, limitations in the procedure for hiding the light source. The result is consistent with an earlier claim that the visibility or otherwise of the light source does not affect the illusion (Schroeder 1852) and, taken at face value, it would appear that neither of these factors had a major effect on the strength of the illusion. The fact that the illusion still works even when a light source that is inconsistent with the illusory percept is clearly visible suggests that perception does not rely on explicit reconstruction of scene and illumination. If it did, we would expect the illusion to be weakened when an incompatible light source is clearly visible.

There was also no difference between left and right lighting. This contrasts with previous evidence for a leftwards bias in the interpretation of pictures when lighting is from above (Mamassian and Goutcher 2001; Sun and Perona 1998; but see McManus et al 2004). This may be because our task and stimuli were very different or because there was no vertical component in our lighting directions.

In summary, this experiment suggests that the visibility or otherwise of the light source is not critical to the illusion and that the results of the previous experiment are not likely to have been because the light used was from the left rather than the right. The following experiments made use of a virtual hollow-face illusion to vary factors that cannot be so easily controlled when using real objects.

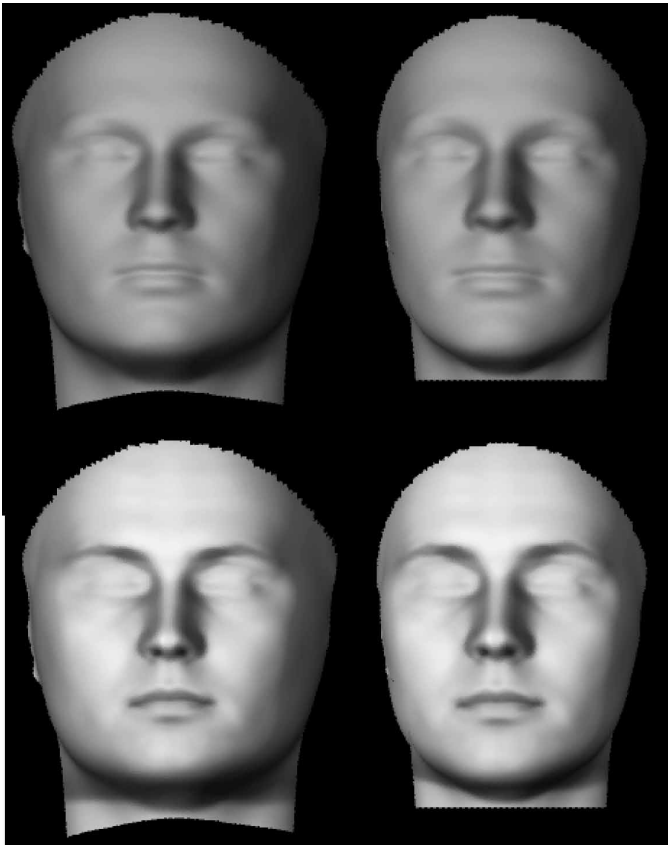
### 5 Experiment 4

The 'virtual' hollow-face illusion used in this and subsequent experiments consists of two rendered rear views of the front half of a three-dimensional head model (Vetter and Troje 1997) presented stereoscopically with the aid of Stereographics CrystalEyes 3 shuttered glasses. Despite the disparity information available to the contrary, the virtual mask, like its real counterpart, appears convex when presented rendered with a uniform grey reflectance or with its associated surface-colour information. The virtual version of the illusion provides a powerful illustration of the combination of information from both eyes without realisation of the implications of the disparities. The perceived image does not correspond to either of the presented images—the nose does not point to either of the eyes but is seen to point straight ahead. Despite this apparent integration of the images, the face is seen as convex rather than concave. As with the physical illusion, when one eye is closed, the nose points in the direction of the eye that remains open, that is in the opposite direction to that for a normal face.

If a random texture is added to the surface of the head model, this provides clear cues to binocular disparity, and the model is veridically perceived as hollow. Adding texture contrast is also known to reduce the precision of curvature from shading information, even when cues are consistent (Johnston and Passmore 1994). The basic method used

in these experiments was to measure how confidence ratings of perceived convexity/concavity varied as a function of the proportion of noise texture under different stimulus conditions. This is similar to varying the contrast of a random-dot slide projected onto a real mask (Georgeson 1979). The advantage of the virtual setup is that it provides ready control over stimulus and presentation properties that are difficult or impossible to control with a real mask. However, screen-based presentation may differ in important ways from real objects.

In experiment 4, we varied a stimulus property, the presence or absence of a surface colour, and a presentation property, whether the projection used was perspective or orthographic (as defined below). In the context of the role of familiarity on the illusion, we wished to test whether addition of a realistic surface colour would affect the strength of the illusion. Most versions of the physical illusion are of uniform reflectance. Although the illusion does still work if features such as lips and eyes are painted appropriately (eg Papathomas and Bono 2004), no one has compared the relative strength of illusions with and without information about surface pigmentation. Informally it appears clear that the coloured face looks more familiar than a grey face (see figure 5) and, if familiarity is important, this might be expected to result in a stronger illusion.



**Figure 5.** Examples of the stimuli used in experiment 4. Left images from each stereo pair are shown, all with 0% noise texture. The top row shows the ‘grey’ condition with Lambertian shading, while the bottom row shows the ‘colour’ condition rendered with a spectral component. The left column is rendered with perspective cues while the right column was rendered in orthographic projection. Although the colour images have higher mean luminance, the image contrasts (standard deviation/mean) were within 2% at all levels of noise. A colour version of this figure is available in the additional materials online (<http://www.perceptionweb.com/misc/p5523/>).

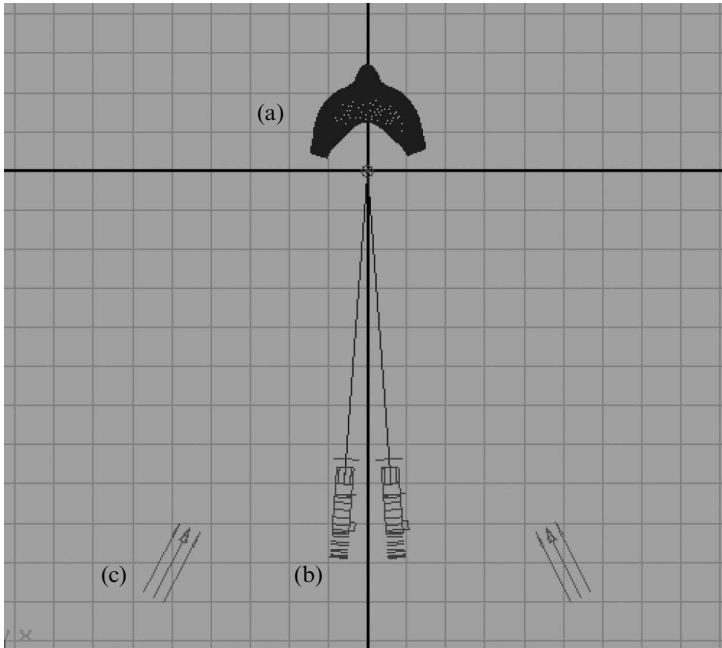
However, there are a number of more objective reasons that a mask rendered with the coloured face might be expected to result in a weaker illusion. First, the surface colour may act like the noise texture in providing high-contrast low-level image features that provide a better input for stereo-disparity processing than low-contrast shading, thereby helping to disambiguate the illusion (Deutch et al 1990). Second, if the ambiguity of shape-from-shading is particularly critical in generating the illusion, a uniform matte albedo might be expected to produce the most effective illusion, since colour may disrupt shape-from-shading, as it is not of itself a cue to shape, and it introduces gradual changes in albedo, which may need to be distinguished from changes in illumination. In particular, the images used to provide colour information are based on video and inevitably contain some shading information related to the original lighting used to record them that is inconsistent with the modeled lighting used for rendering. Third, colour-defined texture elements might also provide a cue to depth when perspective cues are present (van der Enden and Spekreijse 1989). Thus there are at least three properties of the added colour that might be expected to weaken the illusion. Only if colour enhances familiarity or ‘faceness’ might it be expected to enhance the illusion.

The second experimental manipulation used in this experiment was whether images were rendered with an orthographic or a perspective projection. Perspective is an important pictorial cue to depth and reverse perspective can create highly effective illusions of depth reversal (Papathomas 2002; Papathomas and Bono 2004; Wade and Hughes 1999). Texture perspective in particular has been proposed as a reason why pseudoscopically presented images of a face do not reverse—monocular texture perspective will still indicate convexity (van der Enden and Spekreijse 1989). With the virtual hollow face it is possible to render images without perspective cues to depth by using an orthographic projection. This is an orthogonal projection of the mask onto the image plane such that the object’s  $z$ -values within the camera’s coordinate system have no effect on the image. Orthographic projections might be expected to lead to a more effective illusion as they lack an important cue to veridical depth.

In summary, this experiment compared versions of the virtual hollow face that were rendered with information based on surface colour or with a uniform mid-grey Lambertian reflectance. The different versions of the mask were projected in perspective or orthographically.

## 5.1 Methods

**5.1.1 Materials.** The face and surface-colour information used were based on the average of 200 males and females (Vetter and Troje 1997). This was converted to OBJ format, imported into Maya, trimmed and smoothed to give a model with a total of 25101 vertices. The rear of the head was trimmed also removing the ears leaving a model with maximum dimensions of 20 cm  $\times$  15 cm  $\times$  12 cm (height  $\times$  width  $\times$  depth). Virtual cameras were 40 cm in front of the origin and 6 cm apart, a distance approximating the eye separation on the model. The model itself was centred 7.5 cm behind the origin. The relatively short viewing distance was chosen to accentuate perspective cues and to give suitably sized images for screen-based presentation. The cameras converged on the origin, that is in front of the actual mask, but approximately consistent with convergence for an illusory face (Hoffman and Sebald 2005). Perspective cameras had a focal length of 35 mm, while orthographic cameras had an ‘orthographic width’ of 60 cm (a Maya variable determining how much of the scene the camera sees). The depth-of-field setting was off, so that all surfaces were sharply focused. Lighting was from two directional lights, one from the left with a relative intensity of 0.5 and one from the right with a relative intensity of 1. The model was rendered with a mid-grey Lambertian reflectance or with surface reflectance based on average colour information (Vetter and Troje 1997). Examples of the stimuli are shown in figure 5 and the setup used for rendering in figure 6.



**Figure 6.** The setup used for rendering the virtual hollow face used in experiment 4 and subsequent experiments showing (a) head model, (b) cameras, and (c) directional lights.

The head was also rendered with a Perlin ‘volume noise’ texture (threshold 0; amplitude 1; frequency ratio 2; depth max 3; time 0; frequency 10; scale 1; origin 5, 0, 0, and 0 implode) for left and right views in both projections. These noise images were linearly blended with the corresponding colour and grey images in Matlab with proportions of noise from 0% to 100% in steps of 10%.

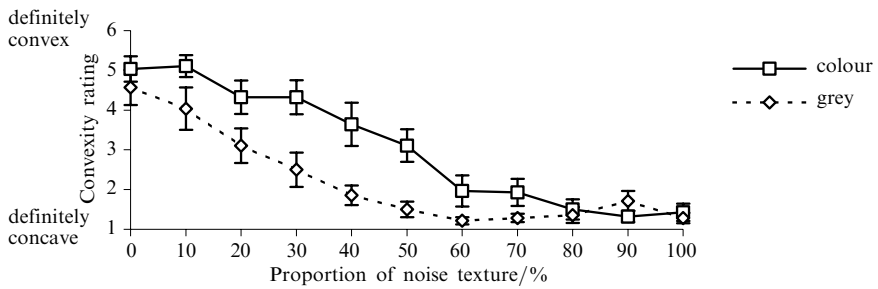
**5.1.2 Design.** The experiment was run as a 2 colour (present, absent)  $\times$  projection (perspective, orthographic)  $\times$  11 noise texture (0%, 10%, ... 100%) full factorial within-subjects design. The dependent variable was the observer’s rating on a six-point scale from 1 (definitely concave) to 6 (definitely convex).

**5.1.3 Procedure.** Observer’s stereoscopic vision was tested with an anaglyph (Frisby 1979). The observers then put on the Stereographics glasses and were shown two practice trials, one a 100% noise texture orthographic grey hollow mask and the other a pseudo-stereoscopically presented (left and right images reversed) colour orthographic projection with 70% noise texture. They were given feedback for practice trials until it was clear that they were seeing the first stimulus as concave and the second as convex. After the practice trials no feedback was given. Each trial was presented in a random order and the stimulus remained up for as long as it took the observer to respond.

**5.1.4 Observers.** Fourteen observers were recruited from local universities in Japan. All had normal or corrected-to-normal vision and were tested for functioning stereopsis as described above.

## 5.2 Results and discussion

A 2 (projection)  $\times$  2 (colour)  $\times$  11 (texture) ANOVA gave a significant colour  $\times$  texture interaction ( $F_{10,130} = 3.9$ ,  $p < 0.05$ ) but no effect of projection or any interactions involving this variable (all  $ps > 0.1$ ). As can be seen from figure 7, convexity ratings decreased with increasing proportion of noise texture. It is also clear that, overall, ratings for coloured stimuli were higher than for grey stimuli, although this difference disappears at both



**Figure 7.** The results of experiment 4. Ratings are shown as a function of surface-colour information and the proportion of noise texture. High ratings indicate the illusory percept. Error bars indicate  $\pm 1$  SEM.

extremes of noise texture, 0% and 100%. Simple-main-effects analysis showed effects of colour for texture levels 2–6 at  $p < 0.05$  and, marginally, for texture levels 7 and 8.

The main finding of this experiment was that rendering the model with natural colour information results in higher convexity ratings. It appears that the increased ‘faceness’ of the stimuli outweighs the additional high-contrast pigment information that might have provided better input for stereopsis. Anecdotally, we have added ‘freckles’ to the colour information but, again, these do not seem sufficient to disambiguate the illusion.

The addition of a random-noise texture does cause the model to be seen veridically as hollow. This appears to be primarily because the additional high-contrast information provides a better input for stereo vision. If texture gradients were critical, we would have expected a difference in the effect of noise texture between perspective and orthographic conditions, as texture gradient cues are absent in the latter. This suggests that the result reported by van der Enden and Spekreijse (1989) was also primarily because of the additional cues to binocular disparity provided in their ‘neutral texture’ condition, rather than the masking of texture perspective cues (Deutch et al 1990).

The absence of an effect of projection confirms that, although perspective can be an effective cue to depth, it is not sufficient to disambiguate the illusion (as is clear from physical masks where this cue is available). Indeed, within the limitations in sensitivity of this method, it does not appear to significantly change the strength of the illusion. One possibility is that faces in reverse perspective are seen as different, fatter (due to the increase in perceived size of the boundary of the face resulting from its proximity) faces rather than as a depth-inverted version of the original face. In this sense, perspective is ambiguous, as in the example of a reversed three-dimensional cube being perceived as a truncated pyramid (Shopland and Gregory 1964). The addition of a texture did not introduce an effect of projection.

In summary, the addition of surface-colour information appears to enhance the hollow-face illusion, consistent with it increasing the likelihood of perceiving the illusion as a normal, convex face. This effect appears stronger than can be achieved by any enhancement of stereo disparity or texture perspective cues associated with the high-contrast boundaries provided by the surface-colour information. Within the limitations of this method, the presence or absence of perspective cues does not affect the strength of the illusion.

## 6 Experiment 5

Using a virtual hollow face also allowed us to compare photographic positive and negative versions of the illusion. Photographic negation is well known to disrupt many aspects of face processing including recognition, although it leaves two-dimensional measures of the image unchanged (Johnston et al 1992). One possible explanation of the effect of negation is that it disrupts the recovery of shape-from-shading information.

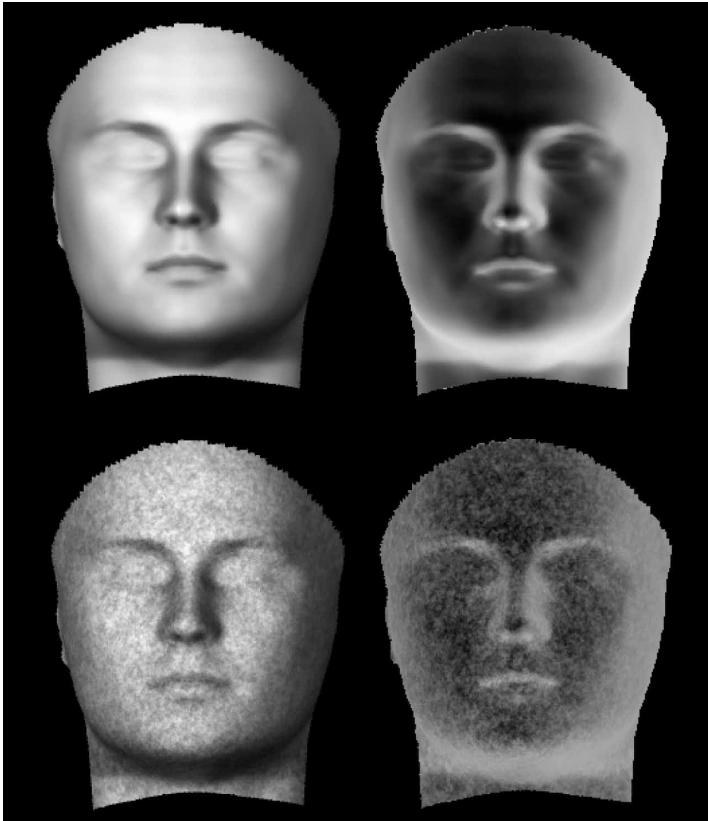
Given that the ambiguity of shading may be central to generating the hollow-face illusion, we decided to compare positive and negative versions of the virtual hollow face, predicting that negative versions would be significantly less effective. Negation also affects the appearance of pigmented areas—in this experiment the pigment information was carried by the RGB images used in experiment 4 converted to grey-scale. Although the eyes are closed, the reversal of the luminance of the lips and eyebrows may also reduce the strength of the illusion by rendering the mask less face-like in appearance. Again, it would be difficult to test the effect of this manipulation with a real mask (although see Anstis 1992).

### 6.1 Methods

Ten observers recruited from local universities in Japan took part in this experiment.

Images were prepared of the hollow face rendered with noise texture and a grey-scale version of the average surface-colour information (grey-scale was used to simplify photographic negation). For the negative condition, only pixels in the area of the face were transformed in order to leave the background the same. Perspective projection was used throughout. Images were blended with different proportions of noise as before. Noise images were not negated. Examples of the stimuli are shown in figure 8.

Presentation and responses were as for experiment 4. The design was a 2 photographic polarity (positive, negative)  $\times$  11 noise texture (0%–100% in steps of 10%) fully factorial within-subjects design. In this experiment five catch trials were included, where the original left and right images were pseudoscopically reversed to give a genuinely convex appearance. The presentation condition for these catch trials was fully randomised.

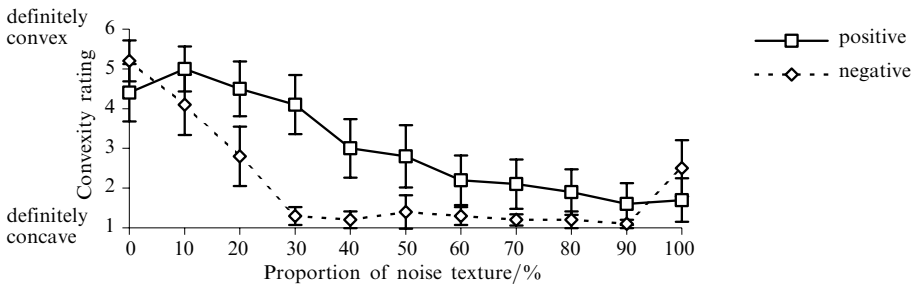


**Figure 8.** Examples of the stimuli used in experiment 5. The left column shows positive images and the right image negatives. The top row has 0% noise texture and the bottom row 60%.



## 6.2 Results and discussion

Analysis of variance gave a polarity  $\times$  texture interaction ( $F_{10,90} = 2.9$ ,  $p \ll 0.05$ ). As can be seen from figure 9, convex ratings fell off more quickly for negative than for positive stimuli. At the  $p < 0.05$  level there were simple main effects of polarity for texture levels 4 and 5 (30% and 40%) and a marginal,  $p < 0.1$ , effect for texture level 6 (50%), with positive images always rated more convex than negative images. The trend was maintained for all except the extreme values of noise texture, 0% and 100%, where positive and negative did not differ significantly. There were simple main effects of texture for both positive and negative images ( $F_{10,90} = 7.1$ ,  $p \ll 0.05$  and  $F_{10,90} = 11.3$ ,  $p \ll 0.05$ , respectively). The catch trials were given an average rating of 5.2 (SEM 0.3), showing that they were correctly perceived as convex.



**Figure 9.** The results of experiment 5. Ratings are shown as a function of image polarity and the proportion of noise texture. High ratings indicate the illusory percept. Error bars indicate  $\pm 1$  SEM.

For low levels of noise texture, even the negative images are perceived as convex. This suggests that, while the ambiguity of shading is important to generating the hollow-face illusion, it is not essential. Hollow surfaces represented as contours or polygonal facets (without shading) are also often depth ambiguous, suggesting that the ambiguity reflects general properties of surface representation. In both these cases, the lines are related to the shape of the object represented. This is also true of the isophotes in negatives, which are the same as for positives, and which therefore have the same geometry. The patterns of isophotes, with additional information about the order in brightness, indicate the position of hills, ridges, and valleys (Koenderink and van Doorn 1980). Even negative faces are recognisable as faces and may tap into the implicit assumption that faces are convex (the single image presented in the negative appears convex, although the pattern of shading is not readily interpretable—see figure 8). The tendency to see negatives as convex soon drops off when veridical stereo cues to depth available from the noise texture are added.

In summary, negating images of the hollow face reduced the tendency for it to appear convex. This could be both because the manipulation interferes with perception of the face as a face and also because it disrupts the perception of shading. Negative versions of the face presented without noise texture added were perceived as convex, suggesting that shading is not essential to the illusion. However, when normal shading cues are available, the illusion persists at higher proportions of added noise texture. Negative faces are like upside-down faces in that, while they still generate an illusion, it is not as strong as for upright positive faces. The effect of negation may also be to disrupt familiar patterns of shading that the results of experiment 2 suggested are an important component of familiarity.

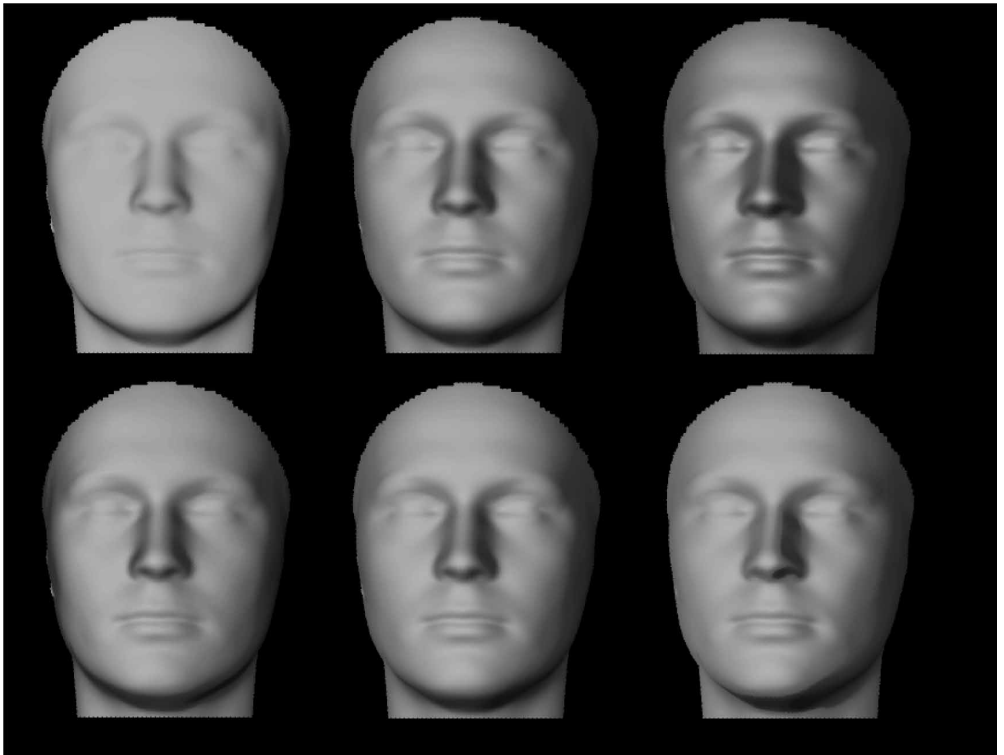
## 7 Experiment 6

One of the original aims of using a virtual hollow face was to investigate the effect of absolute depth on the strength of the illusion. The objects used in the first experiment reported here, like those used in similar experiments (Klopfer 1991), had different

depths and it is possible that this contributes to the perceptual differences reported. However, in neither of the reported experiments does depth predict the pattern of results obtained. In particular, turning masks upside-down has no effect on depth but does affect their perceived convexity and rigid rotation. Observers also often report that the nose, the deepest feature on the mask, is the most difficult to see as concave, despite being associated with the largest disparities (Yellott 1981). In this experiment, we directly tested whether absolute depth is an important determinant of the strength of the illusion.

When looking at a face or mask in the full-face view, height and width are apparent but depth is not directly visible. Greater depth will lead to larger disparities (visible as the degree of apparent rotation in figure 10) and this might be expected to reduce the strength of the illusion. However, greater depth also increases the contrast of the (ambiguous) shading cues (see top row of figure 10) and this might strengthen the illusion. In order to test the effect of depth, we scaled the virtual mask along the  $z$ -dimension, that is perpendicular to the plane of the face, by 50% and by 150%. Pilot experiments had showed no consistent effect of absolute depth (Hill et al 2003). We reasoned this might be because the effects of increased disparities are offset by the increased contrast of shading. In order to control for this, we used histogram equalisation to equate shading contrast in 50% and 150% depth-scaled images to the unscaled, 100%, mask (see bottom row of figure 10).

As noted in the introduction, direct measures of relative depth suggest a flattening of the illusory face (Hartung et al 2005; Kroliczak et al, in press). In this case, the positively depth-scaled face might be perceived as having the appropriate relative depths



**Figure 10.** Examples of the stimuli used in experiment 6. Images from left to right correspond to 50%, 100%, and 150% depth-scaled images. The top row shows unequalised stimuli while the stimuli in the bottom row have been histogram equalised to match the 100% depth images.

when perceived as an illusion. Work with rendered dents has shown that concave surfaces in general appear flatter than their convex equivalents and that, for concave surfaces, judgments are independent of simulated depth (Liu and Todd 2004). We did not attempt to measure relative depth directly here but, if flattening is important, the 150% depth-scaled illusion may prove the most compelling.

In summary, this experiment was designed to test if scaling the depth of the mask and associated binocular disparities are critical to the illusion.

### 7.1 Methods

Ten observers recruited from local universities in Japan took part in this experiment.

Methods were as described previously, except for the following. Shaded images of 50%, 100%, and 150% depth-scaled masks were rendered with Lambertian reflectance or with volume Perlin noise in Maya 5. Noise parameters were threshold = 0; amplitude = 1; ratio = 1; frequency ratio = 2; depth max = 3; frequency = 10; and origin = 20, 0, 0. Noise images were rendered with ambient lighting so as to eliminate shading cues dependent on depth in the noise images. For the shaded face, two directional lights were used at an angle of  $26.6^\circ$  to right and left of the normal to the face with intensities 1 and 0.5, respectively, as for previous experiments. Orthographic projection was used throughout. The virtual cameras were separated by 6 cm, approximately the interocular distance of the head model, positioned at a distance of 40 cm and aimed at what would have been the centre of the head, the distance of the nearest edges of the mask. The maximum depth of the original mask was approximately equivalent to 12.5 cm, meaning that 50% scaling corresponded to a maximum depth of 6.3 cm and 150% to 18.8 cm.

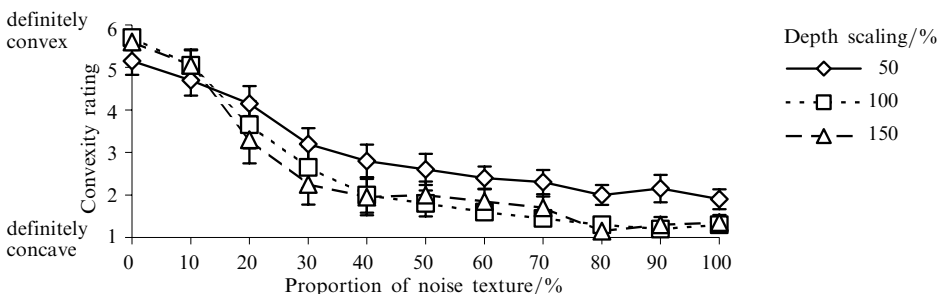
### 7.2 Design

The design was a 3 depth (50%, 100%, 150%)  $\times$  2 contrast (equalised, unequalised)  $\times$  11 texture (0%–100%, in 10% steps) within-subjects fully factorial design. The dependent variable was rated on a scale with 1 indicating definitely concave and 6 definitely convex.

### 7.3 Results and discussion

An ANOVA gave a depth  $\times$  texture interaction ( $F_{20,180} = 2.0$ ,  $p < 0.05$ ). As can be seen from figure 11, the 50% scaled stimuli produced a different pattern of responses from the 100% and 150% versions. Relative to the 100% and 150% stimuli, the 50% stimuli were given higher convexity ratings at high noise texture levels but lower ratings when the proportion of noise texture was low. There was no effect of equalisation on the overall analysis ( $p > 0.1$ ).

As there is no difference between equalised and unequalised images for 100% stimuli, we repeated the analysis omitting these stimuli. This showed an effect of quantisation ( $F_{1,9} = 10.3$ ,  $p < 0.05$ ), with equalised images rated slightly higher than unequalised



**Figure 11.** The results of experiment 6. Ratings are shown as a function of depth scaling and the proportion of noise texture. High ratings indicate the illusory percept. Error bars indicate  $\pm 1$  SEM.

images: mean (SEM) 2.9 (0.2) and 2.7 (0.2), respectively. This effect did not interact with depth ( $p > 0.1$ ), suggesting that increased shading contrast did not necessarily result in higher ratings—contrast is greater when equalised for 50% stimuli but less for 150%. Instead, the ranges of contrast associated with unscaled stimuli may be most effective.

The results suggest that relatively shallow stimuli, like cameos and intaglios, are more ambiguous than deeper stimuli, and are less likely to be rated as either definitely convex or definitely concave. Their actual flattening may also be compounded by illusory depth compression (Hartung et al 2005; Kroliczak et al, in press) further reducing confidence with respect to perceived convexity. However, over a certain limit, depth scaling does not appear to make any difference to ratings. This suggests that differences in depth are likely to have made at most a minor contribution to the differences reported between the objects in experiment 1 and elsewhere (Klopfer 1991), especially given all the objects did not vary to the extent of the masks used here.

In summary, depth range does not appear to be a major determinant of the strength of the illusion, explaining why the nose is not necessarily the easiest feature to be seen veridically. Instead, faces of any depth may be fitted to an internal representation corresponding to the 100% depth, with changes in depth perceived as changes in lighting (see figure 10). Equalising intensity values slightly increases ratings whether it enhances or reduces contrast, but does not offset the effects of disparity.

## 8 General discussion

A series of experiments have been reported that were designed to test whether the hollow-face illusion is a result of object-specific top-down knowledge, general rules used for interpreting bottom-up data, or the bottom-up data themselves. Below we discuss the results reported within the framework of these different classes of explanation.

It was argued in the introduction that primary evidence for the importance of familiarity in the illusion comes from the detrimental effect of turning the hollow face upside down (Hill and Bruce 1993, 1994; Klopfer 1991; Papathomas and Bono 2004). Experiment 1 demonstrated this effect for objects other than the face. The results also suggested that this effect of familiarity is not 'all-or-nothing' but graded, and that the effect is probably strongest for faces.

Experiment 2 was a step towards trying to define what factors contribute to the familiarity of an object, particularly a face. The results showed that the differential familiarity of upright and inverted orientation is not dependent on the light appearing to be from above. The experiment also showed that there was a difference between light from above and below defined with respect to the object. The pattern of shading resulting from light from above the forehead is more familiar than that associated with light from below the chin and, like the differential familiarity between upright and inverted orientations, this results in a stronger illusion. Thus the large effect of inversion found for the hollow face when it is rotated under natural lighting conditions has at least two components: the familiar configuration of features and the familiar pattern of shading.

Another possible component of familiarity was found in experiment 4, the pattern of surface-colour information. These pigment differences might have been expected to facilitate bottom-up cues to concavity including texture perspective and stereo. However, the addition of a familiar texture map was found to strengthen the illusion, emphasising the importance of familiarity relative to the additional bottom-up cues.

In summary, although even unfamiliar concave objects can generate the illusion of convexity (Deutch et al 1990; Hill and Bruce 1994; Mach 1886/1959), relative familiarity can contribute considerably to the strength of the illusion for the face and other objects. Familiarity reflects prior experience with objects such as faces and is associated with particular orientations and patterns of lighting and pigmentation. The effect

of familiarity presumably reflects access to the same stored representations as are used for recognition and other tasks. When these are consistent with at least some of the bottom-up data, in particular 2-D distributions of features and the ambiguous shading, all or most of the data are perceived in a way that fits with these existing representations. The results suggest that such internal representations are not perfect object-centred models of the kind proposed by Marr (Marr and Nishihara 1978), but reflect prior experience. The existence of the illusion in itself suggests that 3-D depth is an important component of these representations.

The second class of explanations outlined in the introduction was based on knowledge about the world embodied as general rules or assumptions. Experiment 2 sought to test if explicit knowledge of the light-source position combines with knowledge of shadows and shading in the generation of the illusion (Rittenhouse 1786; Wade 1983). However, the visibility or otherwise of the light source was not found to affect the illusion. The fact that the illusion still works even when an inconsistent light source is clearly visible highlights a disjunction between our explicit knowledge and perception. It also suggests that a full reconstruction of the scene including illumination is not necessary and that our perception tolerates local anomalies. This is also true of perception of the illusion under natural viewing conditions, for example outside on a sunny day. In this case, the perceived lighting direction for the illusion is inconsistent with that for the rest of the (vertically perceived) objects in the scene. This violates another proposed assumption, a shared single light source for the objects in the scene (Ramachandran 1988). Informal experimentation with the illusion under various conditions cannot help but suggest that lighting is critical, and the illusion can be disrupted or enhanced by the choice of lighting conditions, surface properties, and the associated specularities. However, the results of experiment 2 and 3 suggest that the effect of lighting is based on the object-centred pattern of shading rather than on explicit or implicit reconstruction of the scene, including light sources.

In experiment 5 photographic negation was used to disrupt shading. The advantage found for positive over negative is consistent with the importance of shading to the illusion. However, that the illusion works with negative (or contour defined) versions shows that a familiar pattern of shading is not essential for the illusion. The perception of negative stereo pairs as convex may reflect a general preference for convexity. However, this cannot constitute a complete explanation as we are perfectly capable of seeing concavities. The question remains why the convex interpretation should be favoured under certain circumstances (Yellott 1981). In the case of negatives, the isophotes and other 2-D properties are unchanged from the positive and these may be enough to establish associations between the stimulus and the familiar, convex, face. If this were the case, we would expect an additional effect of turning the negative upside down. For unfamiliar objects, factors that contribute to them being seen as lumps of matter, and therefore globally convex, rather than concave features on larger objects, may have a role to play and help to understand scene segmentation.

The last class of explanations includes those based solely on bottom-up stimulus properties. The hollow face, ordinary paintings, and reverspectives all demonstrate that many bottom-up cues are ambiguous. This is critical to the illusion. When stereo disparity cues alone are provided, as when the virtual mask is presented rendered with a high-contrast noise texture, the face is no longer perceived as convex (Georgeson 1979). Disparities are unambiguous, and the familiarity of faces as convex objects by itself is not enough to generate the illusion. Some bottom-up data consistent with the convex solution, even if ambiguous, must be present. The importance of lighting and photographic polarity suggests that shading may be particularly important in the context of the face (Bruce 1988), but other ambiguous bottom-up cues including contours can also be sufficient.

Any deterministic account solely in terms of bottom–up properties must address a number of unanswered questions, primarily what happens to the unambiguous stereo information that is available. As outlined in the introduction, previous evidence shows that the sign of the disparity is not mentally reversed (Georgeson 1979; Yellott and Kaiwi 1979). Nor is stereo information simply vetoed, as clearly demonstrated by the virtual hollow face with its nose point directly out of the screen and not towards either particular eye. Like the motion-parallax information, stereo must be ‘explained away’ in some way. However, unlike motion parallax, which is reinterpretable as rotation, there is no clear alternative interpretation available for binocular disparities. Instead, it seems that some, but not all, of the implications of the disparities are accepted—disparities still signal that the surface is 3-D, but the depth implications of the sign of disparity are ignored. One possibility would be that sign is reversed at a relatively late stage, for example at the level of surface representation in terms of curvature rather than depth itself (Koenderink 1990).

The effects of absolute depth reported in experiment 6 and previously (Hill et al 2003) are somewhat unclear. Increasing depth had no effect but decreasing depth did appear to make the stimuli more ambiguous, with observers giving more intermediate ratings. Equalising shading contrast did not appear to modify the effect of depth, although equalised shading was given higher ratings irrespective of depth. There was not a deterministic relationship between the apparent depth and the physical magnitude of the disparities, suggesting ‘fusion without stereopsis’ (Yellott and Kaiwi 1979) rather than a change in sign. This would be consistent with incoming sensory data being fit to existing representations with their associated depths and the magnitude of disparities largely ignored. This is one experiment where, through the use of a 3-D printer, it would be possible to compare real and virtual presentation to determine if results are consistent.

The effect of inversion also poses a problem for bottom–up accounts of the illusion (Hill and Bruce 1993, 1994; Klopfer 1991; Papathomas and Bono 2004). With inversion the object acts as its own control with regards to bottom–up properties including depth, and yet orientation appears to have a large effect on the illusion. If these effects were due to chance bottom–up features, for example particularly prominent specularities visible in one orientation but not the other, there is no reason to expect the stronger illusion to be consistently associated with the familiar, upright, orientation. A bottom–up account also needs to explain why the addition of surface-colour information strengthens the illusion despite providing additional disambiguating bottom–up cues and also why the projection used has little or no effect despite potentially affecting many of these cues.

In summary, the hollow-face illusion appears to reflect a combination of the explanations offered. Some ambiguously interpretable bottom–up data must be present. That, coupled with a general bias towards convexity, is sufficient to generate the illusion, even when this interpretation is incompatible with other, unambiguous, bottom–up data. However, familiar orientations and patterns of shading and surface-colour information can greatly enhance this effect for both faces and other familiar objects.

**Acknowledgments.** This research was conducted as part of ‘Research on Human Communication’ with funding from the National Institute of Information and Communications Technology. Thanks to everyone who participated in the experiments and provided stimulating discussions and especially to A Chakraborty, J Bennett, V Cockram, and E Moore who ran experiment 1 as their miniproject, Roger Gassert for translating Schröder, Glyn Cowe for developing the stereographic display software used to present the virtual hollow face, and two anonymous reviewers for their constructive criticism.

---

**References**

- Anstis S M, 1992 "Visual adaptation to a negative, brightness-reversed world: Some preliminary observations", in *Neural Networks for Vision and Image Processing* Eds G A Carpenter, S Grossberg (Cambridge, MA: MIT Press) pp 1–14
- Belhumeur P N, Yuille A, Epstein R, 1996 "Learning and recognizing objects using illumination subspaces", in *International Workshop on Object Representation for Computer Vision*, Cambridge (Berlin: Springer)
- Berbaum K, Bever T, Chung C S, 1983 "Light source position in the perception of object shape" *Perception* **12** 411–416
- Blake A, Bülthoff H H, 1990 "Does the brain know the physics of specular reflection?" *Nature* **343** 165–168
- Brewster D, 1826 "On the optical illusion of the conversion of cameos into intaglios, and of intaglios into cameos, with an account of other analogous phenomena" *Edinburgh Journal of Science* **4** 99–108
- Brewster D, 1832 *Letters on Natural Magic* Addressed to Sir Walter Scott, Bart (London: John Murray)
- Bruce V, 1988 *Recognizing Faces* (London: Lawrence Erlbaum Associates)
- Deutch J A, Ramachandran V S, Peli E, 1990 "Binocular depth reversals despite familiarity cues: An artifact?" *Science* **279** 565
- Emrich H, 1989 "A three-component-system hypothesis of psychosis. Impairment of binocular depth inversion as an indicator of functional equilibrium" *British Journal of Psychiatry* **155** 37–39
- Enden A van der, Spekrijse H, 1989 "Binocular depth reversal despite familiarity cues" *Science* **244** 959–961
- Frisby J, 1979 *Seeing: Illusion, Brain, and Mind* (Oxford: Oxford University Press)
- Georgeson M A, 1979 "Random-dot stereograms of real objects: observations on stereo faces and moulds" *Perception* **8** 585–588
- Gilchrist A, 2003 "The importance of errors in perception", in *Colour Perception: Mind and the Physical World* Ed. R M D Heyer (Oxford: Oxford University Press)
- Gmelin P F, 1744 "Epistola de radice ipecacuanhae observationes quasdam medico-physicas & de fallaci visione per microscopia" *Philosophical Transactions of the Royal Society* **43** (reprinted edition) 382–391
- Gregory R L, 1970 *The Intelligent Eye* (London: Weidenfeld and Nicolson)
- Gregory R L, 1973 "The confounded eye", in *Illusion in Nature and Art* Eds R L Gregory, C E Gombrich (London: Duckworth) pp 49–95
- Gregory R L, 1980 "Perceptions as hypotheses" *Philosophical Transactions of the Royal Society of London, Series B* **290** 181–197
- Gregory R L, 1997 "Knowledge in perception and illusion" *Philosophical Transactions of the Royal Society of London, Series B* **352** 1121–1128
- Hartung B, Schrater P R, Bülthoff H H, Kersten D, Franz V H, 2005 "Is prior knowledge of geometry used in visually guided reaching?" *Journal of Vision* **5** 504–514
- Heard P F, Chugg C, 2003 "Top-down knowledge affects the perception of inside-out objects" *Perception* **32** Supplement, 50 [abstract]
- Hill H, Bruce V, 1993 "Independent effects of lighting, orientation, and stereopsis on the hollow-face illusion" *Perception* **22** 887–897
- Hill H, Bruce V, 1994 "A comparison between the hollow-face and 'hollow-potato' illusions" *Perception* **23** 1335–1337
- Hill H, Chakraborty A, Evans H, Cowe G, Johnston A, 2003 "The hollow-face and related illusions: object knowledge or geometric properties", Technical Report of the Institute of Electronics Information and Communication Engineers, HIP2003-88, pp 61–66
- Hoffman J, Sebald A, 2005 "Does the vergence angle adapt to the hollow face illusion?", in *Proceedings of the 14th Meeting of the European Society for Cognitive Psychology (ESCoPXIV)* (Leiden: Department of Cognitive Psychology, Leiden University) p.77; available online at [http://www.psychologie.leidenuniv.nl/cog/content\\_docs/ESCoP2005/proceedingsweb2.pdf](http://www.psychologie.leidenuniv.nl/cog/content_docs/ESCoP2005/proceedingsweb2.pdf)
- Horn B K P, 1977 "Understanding image intensities" *Artificial Intelligence* **8** 201–231
- Howard I P, 1990 "Shape from shading in different frames of reference" *Perception* **19** 523–530
- Howard I P, Rogers B, 1995 *Binocular Vision and Stereopsis* (New York: Oxford University Press)
- Howard I P, Rogers B, 2002 *Seeing in Depth* (Thornhill, Canada: I. Porteous)
- Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365–375

- Johnston A, Passmore P, 1994 "Shape-from-shading I: Surface curvature and orientation" *Perception* **23** 169–189
- Kersten D, Mamassian P, Yuille A, 2004 "Object perception as Bayesian inference" *Annual Review of Psychology* **55** 271–304
- Klopfer D S, 1991 "Apparent reversals of a rotating mask: A new demonstration of cognition in perception" *Perception & Psychophysics* **49** 522–530
- Koenderink J J, 1990 *Solid Shape* (Cambridge, MA: MIT Press)
- Koenderink J J, Doorn A J van, 1980 "Photometric invariants related to solid shape" *Optica Acta* **27** 981–996
- Kroliczak G, Heard P F, Goodale M A, Gregory R L, 2006 "Dissociation of perception and action unmasked by the hollow-face illusion" *Brain Research* **1080** 9–16
- Leweke F M, Schneider U, Radwan M, Schmidt E, Emrich H M, 2000 "Different effects of nabilone and cannabidol on binocular depth inversion in man" *Pharmacology, Biochemistry and Behaviour* **66** 175–181
- Leweke F M, Schneider U, Thies M, Muente T F, Emrich H M, 1999 "Effects of synthetic Delta90tetrahydrocannabinol on binocular depth inversion of natural and artificial objects in man" *Psychopharmacology* **142** 230–235
- Liu B, Todd J T, 2004 "Perceptual biases in the interpretation of 3D shape from shading" *Vision Research* **44** 2135–2145
- Mach E, 1886/1959 *The Analysis of Sensation* (New York: Dover)
- McManus I, Buckman J, Woolley E, 2004 "Is light in pictures presumed to come from the left side?" *Perception* **33** 1421–1436
- Mamassian P, Goutcher R, 2001 "Prior knowledge on the illumination position" *Cognition* **81** B1–B9
- Marr D, Nishihara H K, 1978 "Representation and recognition of the spatial organization of three-dimensional shapes" *Proceedings of the Royal Society of London, Series B* **200** 269–294
- Moore C, Cavanagh P, 1998 "Recovery of 3D volume from 2-tone images of novel objects" *Cognition* **67** 45–71
- Papathomas T V, 2002 "Experiments on the role of painted cues in Hughes's reverspectives" *Perception* **31** 521–530
- Papathomas T V, Bono L M, 2004 "Experiments with a hollow mask and a reverspective: Top–down influences in the inversion effect for 3-D stimuli" *Perception* **33** 1129–1138
- Perrett D I, Oram M W, 1998 "Visual recognition based on temporal cortex cells: viewer-centred processing of pattern configuration" *Zeitschrift für Naturforschung C* **53** 518–541
- Ramachandran V S, 1988 "Perception of shape-from-shading" *Nature* **331** 163–166
- Ramachandran V S, 1995 "2-D or not 2-D?—that is the question", in *The Artful Eye* Eds R Gregory, J Harris, P Heard, D Rose (Oxford: Oxford University Press) pp 249–267
- Rittenhouse D, 1786 "Explanation of an optical deception" *Transactions of the American Philosophical Society* **2** 37–43
- Rogers B, 2003 "Seeing depth (Okuyuki o miru)", in *Nihon Shinrigakukai* (Tokyo: Tokyo University)
- Sakurai K, Endo M, Maruyama K, 1985 "Depth reversal in aperture viewing of facial shading patterns" *Tohoku Psychological Folia* **44** 70–80
- Schneider U, Borsutzky M, Seifert J, Leweke F M, Huber T J, Rollnik J D, Emrich H M, 2002 "Reduced binocular depth inversion in schizophrenic patients" *Schizophrenia Research* **53** 101–108
- Schroeder H, 1852 "Ueber eine optische Inversion mit freiem Auge" *Annalen der Physik und Chemie* **163** 306–312
- Shashua A, 1997 "On photometric issues in 3D visual recognition from a single 2D image" *International Journal on Computer Vision* **21** 99–122
- Shopland G, Gregory R L, 1964 "The effect of touch on a visually ambiguous three-dimensional figure" *Quarterly Journal of Experimental Psychology* **16** 66–70
- Sun J, Perona P, 1998 "Where is the sun?" *Nature Neuroscience* **1** 183–184
- Tarr M J, Kersten D, Bühlhoff H H, 1998 "Why the visual recognition system might encode the effects of illumination" *Vision Research* **38** 2259–2275
- Valentine T, 1988 "Upside-down faces: a review of the effect of inversion upon face recognition" *British Journal of Psychology* **79** 471–491
- Vetter T, Troje N F, 1997 "Separation of texture and shape in images of faces for image coding and synthesis" *Journal of the Optical Society of America A* **14** 2152–2161
- Wade N J, 1983 *Brewster and Wheatstone on Vision* (London: Academic Press)
- Wade N J, Hughes P, 1999 "Fooling the eyes: *trompe l'oeil* and reverse perspective" *Perception* **28** 1115–1119



- 
- Wheatstone C, 1838 “Contributions to the physiology of vision—Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision” *Philosophical Transactions of the Royal Society of London* **128** 371–394
- Wheatstone C, 1852 “The Bakerian Lecture—Contributions to the physiology of vision—Part the second. On some remarkable, and hitherto unobserved, phenomena of binocular vision (continued)” *Philosophical Transactions of the Royal Society of London* **142** 1–17
- Wolf R, Wolf D, 1990 “Binocular depth reversal is controlled by censorship process based on familiarity cues”, in *Proceedings of the 6th Annual Meeting of the International Society for Psychophysics* Ed. F Muller (Würzburg: International Society for Psychophysics)
- Yellott J I Jr, 1981 “Binocular depth inversion” *Scientific American* **245**(1) 118–125
- Yellott J I Jr, Kaiwi J L, 1979 “Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse perspective” *Perception* **8** 135–142

ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

# PERCEPTION

VOLUME 36 2007

[www.perceptionweb.com](http://www.perceptionweb.com)

**Conditions of use.** This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.