Perception, 2011, volume 40, pages 975 – 988

doi:10.1068/p6915

Binocular disparity magnitude affects perceived depth magnitude despite inversion of depth order

Harold Matthews, Harold Hill, Stephen Palmisano
School of Psychology, University of Wollongong, Wollongong, NSW 2522, Australia;
e-mail: hm436@uowmail.edu.au
Received 24 January 2011, in revised form 8 July 2011; published online 22 September 2011

Abstract. The hollow-face illusion involves a misperception of depth order: our perception follows our top – down knowledge that faces are convex, even though bottom – up depth information reflects the actual concave surface structure. While pictorial cues can be ambiguous, stereopsis should unambiguously indicate the actual depth order. We used computer-generated stereo images to investigate how, if at all, the sign and magnitude of binocular disparities affect the perceived depth of the illusory convex face. In experiment 1 participants adjusted the disparity of a convex comparison face until it matched a reference face. The reference face was either convex or hollow and had binocular disparities consistent with an average face or had disparities exaggerated, consistent with a face stretched in depth. We observed that apparent depth increased with disparity magnitude, even when the hollow faces were seen as convex (ie when perceived depth order was inconsistent with disparity sign). As expected, concave faces appeared flatter than convex faces, suggesting that disparity sign also affects perceived depth. In experiment 2, participants were presented with pairs of real and illusory convex faces. In each case, their task was to judge which of the two stimuli appeared to have the greater depth. Hollow faces with exaggerated disparities were again perceived as deeper.

1 Introduction

The hollow-face illusion is the perception of a convex face with a nose pointing towards you when viewing the concave side of a cast or a mask of a face (Gregory 1973). The illusion can be characterised as a mistaken perception of depth order, eg the tip of the nose appears closer than all other points although it is actually the farthest away. The illusion is thought to be the result of prior top – down knowledge of faces as typically convex objects dominating bottom – up sensory data (Gregory 1980). When viewed binocularly, the disparity between the left-eye and right-eye views (binocular disparity) provides an unambiguous cue to the actual depth order (it signals that the surface is concave). We perceive the illusory nose as pointing toward our 'cyclopean' eye, demonstrating that left and right images are fused and stereoscopic depth information is available. This raises the question: what happens to the information available from binocular disparities (Hill and Johnston 2007; Yellott 1981; Yellott and Kaiwi 1979). Binocular disparities have both sign (indicating depth order) and magnitude (indicating depth magnitude), and we address this question by asking whether disparity magnitude still affects the perceived depth of the illusory convex face, despite this percept being inconsistent with the depth-order information conveyed by disparity sign.

One possibility is that input from one eye is suppressed (Wheatstone 1852; Yellott 1981). However, as stated above, we perceive the illusory nose as pointing towards a central, 'cyclopean', eye (Hill and Johnston 2007; Yellott and Kaiwi 1979). This effect is only possible if images from the left and right eyes are being combined. However, this combination may involve the "vetoing" of depth information from disparity (Hartung et al 2005), thus constituting "fusion without stereopsis" (Yellott and Kaiwi 1979). The operation of a mechanism that vetoes inconsistent information has been demonstrated in cue conflict studies (Bülthoff and Mallot 1988; Norman and Todd 1995).

Norman and Todd, for example, showed displays where information from different depth cues implied different surfaces. Binocular disparity indicated either vertical or horizontal corrugations, while motion cues indicated corrugations orthogonal to those indicated by disparity. They found that the surface indicated by one of the cues would determine perception, with the implications of the other cue being effectively disregarded. It is possible that a similar mechanism may operate when the conflict is between binocular disparity and prior knowledge. This 'veto' hypothesis is supported by Hill and Johnston's (2007) observation that manipulating the disparity magnitude of the hollow-face did not affect the strength of the illusion.[1] It is worth noting, however, that illusion strength is not necessarily related to perceived depth.

It is only depth-order information from the *sign* of binocular disparity that unambiguously indicates the surface is concave. Another possibility is that this disparity sign information is disregarded but disparity magnitude still affects perceived depth magnitude of the hollow face. This 'sign veto' hypothesis is analogous to a process of "mental pseudoscopy", where left and right images are reversed in the brain (Yellott 1981; Yellott and Kaiwi 1979). "Pseudoscopic reversal" effectively reverses the sign of binocular disparity so that it is consistent with a convex surface (Wheatstone 1852). As such reversal does not affect disparity magnitude this would still be expected to affect the perceived depth magnitude of the illusory face. Evidence against mental pseudoscopy per se has, however, been presented in two compelling demonstrations where a hollow face is perceived incorrectly in reversed depth at the same time as other aspects of perceived depth remain consistent with the sign of binocular disparities present. In one, the depth of a random-dot stereogram projected onto a hollow face is still perceived to have the depth order implied by disparity sign. In the other, a hollow face attached to a Pulfrich pendulum setup is still seen to rotate in the direction expected from the difference in signal timings between the two eyes (Yellott 1981; Yellott and Kaiwi 1979). However, it is still possible that disparity magnitude may affect the depth of the illusory face via some other mechanism than complete mental reversal of left and right images.

A third possibility is that binocular disparity is treated as a unified piece of data, indicating a concave surface with a particular depth magnitude. This is then 'traded-off' against other sources of depth information and the prior knowledge that the surface is convex (Hartung et al 2005). The resulting percept is then a weighted combination of depth estimates, which appears somewhere between an ordinary convex face and the concave face indicated by binocular disparity. Evidence consistent with this possibility comes from the observation that the illusory convex face tends to appear flatter than an equivalent convex face (Hartung et al 2005; Kroliczak et al 2006). That binocular disparity may be traded-off against prior knowledge is also supported by evidence that the hollow-face illusion is weaker when viewed binocularly rather than monocularly (Hill and Bruce 1993; Hill and Johnston 2007; Papathomas and Bono 2004). Increasing the availability of high-contrast image features, conveying information about binocular disparities, by adding a random texture to the hollow-face also increases the probability that the face will appear concave (Georgeson 1979; Hill and Johnston 2007).

In the experiments reported here, we test between these alternative possibilities by manipulating disparity sign and magnitude independently. Sign was manipulated by physically reversing left and right images—a process that leaves all other depth cues unchanged. We manipulated disparity magnitude by varying the intercamera distance

[1] Illusion strength has been quantified as the physical distance from the observer to the hollow face at which perception of the stimulus 'flips' from convex to concave (Hill and Bruce 1993; Hill and Johnston 2007, experiments 1 – 3; Papathomas and Bono 2004, experiment 1), the percentage of time the illusory convex percept dominates the veridical concave one (Papathomas and Bono 2004, experiment 2), and the amount of random texture added to the stimulus before it appears concave (Georgeson 1979; Hill and Johnston 2007, experiments 4 – 6).

used to generate left and right images. This is equivalent to manipulating interocular separation to alter perceptions of depth magnitude, a strategy which is the basis of the Hyperscope (Walker 1986). In experiment 1 participants adjusted the disparity magnitude and, thus, the implied depth from the tip of the nose to the centre of the head of a convex comparison face until its apparent depth matched that of a reference face. A convex comparison face was used as opposed to a point depth-probe as it was the most natural stimulus. In general, 3-D surfaces defined by disparity gradients and curvature may be interpreted differently to point disparities. This is being investigated in ongoing work. Initial results suggest that, in this case, using a depth probe for comparison does not produce very different results from those reported here. The maximum disparity between the convex comparison face images at the point of subjectively equal depth is used as an index of the apparent depth of the reference face. Four reference faces were used which were either convex or concave and with either unexaggerated disparity (disparities were consistent with a normally proportioned face), or exaggerated disparity (consistent with a face stretched in depth). The 'veto' hypothesis predicts that increasing reference-face disparity magnitude will increase perceived depth of convex, but not concave, faces. By contrast, the 'sign veto' hypothesis predicts that increasing reference-face disparity magnitude will increase perceived depth for both convex and concave faces. The 'trade-off' hypothesis predicts that this trend will be the reverse for the hollow faces. In experiment 2 we used the method of constant stimuli, and participants judged which one of a hollow reference face and a convex comparison face appeared deeper. The reference face again had either normal or exaggerated disparities, while the disparity of convex comparison faces varied over a wide range. In this case the proportion of trials in which the illusory face was judged deeper was used as an index of its apparent depth. Again, the 'veto' hypothesis predicts no effect of reference-face disparity, while the 'sign veto' hypothesis predicts that exaggerated disparity hollow faces will appear deeper and the 'trade-off' hypothesis the opposite.

Given the influence of prior top–down knowledge on the hollow-face illusion we also manipulated its influence to observe any effect upon perceived depth. In experiment 1 we manipulated surface colour (in one condition, features were realistically coloured; in the other a uniform mid-grey reflectance was modeled) and in experiment 2 we varied figural orientation (face inverted or not). Both manipulations are known to affect illusion strength by altering (increasing and decreasing, respectively) the familiarity of the stimulus (Hill and Bruce 1993, 1994; Hill and Johnston 2007; Papathomas and Bono 2004).

2 Experiment 1

The aim of experiment 1 was to determine how, if at all, disparity magnitude affects the perceived depth of an illusory convex face. Participants adjusted the depth of a convex comparison face until it appeared to be of the same depth as a reference face (see figure 1). The point at which the depth of reference and comparison faces was judged to be the same—the point of subjective equality (PSE)—was specified in terms of the maximum angular disparity between images (the disparity between the tip of the nose and the simulated fixation point at the base of the mask) generating the convex comparison face at that point. The method of adjustment was employed as it enabled us to compare a large number of conditions without a large number of trials.

The reference face was either convex or concave and maximum disparity magnitude was either $|4.58'|$ or $|13.73'|$ consistent with a face of unexaggerated or exaggerated depth (respectively) under our viewing conditions.[2]

We also manipulated surface colouring of the 3-D faces (either mid-grey or realistic) to determine whether this affected perceived depth magnitude.

[2] Prime ($'$) denotes min of arc.

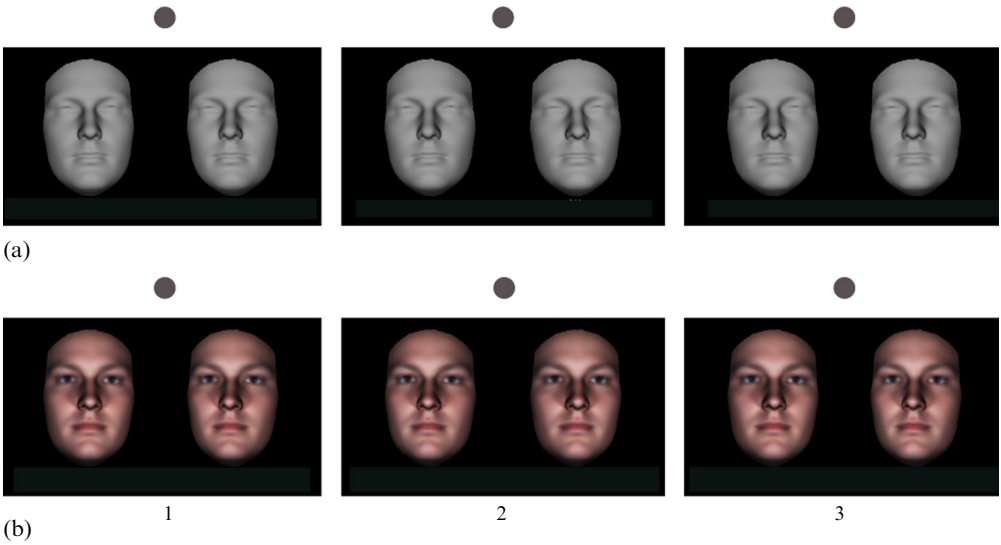

Figure 1. [In colour online, see http://dx.doi.org/10.1068/p6915] When free uncross (1 and 2) or cross (2 and 3) fused this shows two three-dimensional faces like those viewed by participants in the mid-grey (a) and realistic (b) reflectance condition. The face on the left side is a scaled image of a $+9.17'$ comparison face. The face on the right is a $-13.73'$ reference face. Participants adjusted the disparity, and thus implied depth, of the convex comparison face until it appeared to be of the same depth as the illusory convex percept generated by viewing the reference face. There were four reference faces, convex and concave, with either unexaggerated or exaggerated depth.

2.1 *Method*

2.1.1 *Apparatus.* A Wheatstone reflecting stereoscope was used to allow controlled binocular presentation of stimuli. The stereoscope was set up in a black-painted room from which all exterior light sources were blocked during testing. The setup consisted of a chin-rest clamped to the centre of the long side of a table; two 80 mm × 80 mm front-silvered mirrors positioned at eye level 100 mm in front of the axis of the chin-rest; and two Dell 1905FP LCD monitors placed at a distance of 1400 mm to the left and right of the respective mirrors so that the total viewing distance was 1500 mm. Monitor brightness and contrast settings were equated by eye. Black cardboard blinkers were attached to the left and right side of the chin-rest frame to block direct illumination from the monitors in participants' peripheral vision. The two monitors were connected to a PC computer using a 256 Mb ATI Radeon X1300 PRO video card. Monitor refresh rate was set at 60 Hz.

2.1.2 *Stimuli.* The 3-D face model used to create stereo images was based upon an average of 112 male and 112 female faces aged between 19 and 26 years. The face structure and colour information was converted into OBJ format and imported into an empty Blender™ (http://www.blender.org) 3-D environment. The head was cropped so as to remove the ears leaving a model consisting of 7235 vertices with maximum dimensions 128 mm × 193 mm × 66 mm (width × height × depth). The rear of the cropped head was positioned parallel to the $x-y$ plane and centred at the origin of 3-D space. A Lambertian surface reflectance model was applied to the face. The specular component was set to zero. The face was illuminated by a unidirectional 'sun' lighting source with energy set in Blender™ to 0.5. Light was directed parallel to the z axis directly onto the facial surface to avoid producing cast shadows and because lighting from above or below can affect the strength of the illusion (Hill and Bruce 1993, 1996).

Stereo pairs of images were created so that, when displayed on the monitors, they had maximum angular disparity ranging from $0'$ to $|24.78'|$ in increments of $0.91'$. These were taken from two positions to the left and right of the z axis at the same level as head ($y = 0$). We manipulated the maximum disparity between the images by altering the distance between the camera positions while keeping them symmetric about the z axis at a distance of 750 mm. An advantage of this method over scaling the face in depth is that it keeps shading contrast constant. The cameras converged on the centre of the head. When these image pairs were viewed stereoscopically, the implied depth between this simulated fixation point and the tip of the nose increased as a linear function of the maximum disparity between stereo images.

Images were rendered by orthographic projection to remove linear perspective as a cue to depth and ensure convex and concave faces would be equivalent. Ortho-scaling was set to 2.5. The rendered images were 600×600 pixels and consisted of a facial image on a black background. Two sets of images were generated with either realistic colour information or a mid-grey reflectance.

In order to produce the impression of two 3-D faces side by side, stereo pairs of slide shows were created where each slide contained two face images. The image on the left and right side of the slide was horizontally positioned in PowerPoint—156 mm and 113 mm from the top left corner of the slide, respectively. Thus, a portion of the background of the left image was excluded and a portion of the backgrounds overlapped. Both facial images were equidistant from the centre of the slide. Both were vertically positioned 185 mm from the top of the slide. Size was set at 75% for both height and width. Monitor size was 363 mm $\times$ 493 mm (width $\times$ height). Monitor resolution was 1280 by 1240 pixels; thus, facial image size (excluding background) on the monitors was 135 mm and 90.66 mm at its longest and widest points, respectively.

The stereo-pairs simulated two faces: one was a reference face which remained constant during the trial and the other was a comparison face that the participant adjusted. There were 4 reference faces with maximum angular disparity $\pm 4.58'$ or $\pm 13.73'$ (positive and negative signs denote that disparity implied the face was convex or concave, respectively). Hollow faces were generated by left/right reversal of their convex counterparts, so that all that differed between convex and concave reference faces was disparity sign. The comparison face was always convex, with disparity varying from $0'$ to $+24.78'$ in increments of $0.91'$. The depth of the comparison face could be adjusted linearly by scrolling forwards or backwards through the PowerPoint slide show. Left and right slide shows were synchronised with PowerShow™ (http://officeone.mvps.org/powershow/powershow.html), a plug-in for Microsoft PowerPoint™.

Between each face slide was a black slide with a white 'X' located in the position where the comparison face would appear. This indicated to participants which of the two faces they should be adjusting.

Whether the reference face appeared on the left or the right was counterbalanced within subjects. Slideshows were duplicated, one with mid-grey stimuli the other with realistic colouring.

2.1.3 *Participants.* Participants were five male and seventeen female undergraduate psychology students from the University of Wollongong aged between 18 and 34 years ($M = 20.50$ years, SD $= 3.44$ years). All reported 20/40 or better vision. They were randomly assigned to either the realistic or mid-grey reflectance conditions (eleven to each). Participants received course credit for participating.

2.1.4 *Procedure.* The study was approved by the University of Wollongong Human Research Ethics Committee (HE 10/071). Before commencing the experiment the demands of the task were outlined and written consent was obtained. Participants were not informed that faces would sometimes be hollow (to minimise the effects of participant expectations).

Each participant was seated in front of the table with his/her head in the chin-rest. The rotation of the mirrors was adjusted so that black and white grids projected from the left and right monitors both appeared to overlay a grid pinned to the back wall directly behind the mirrors. This calibration was to ensure corresponding points on both monitors projected to corresponding points in the participants' left and right eyes. After this, all light sources, apart from the monitors, were turned off.

To familiarise themselves with the setup participants initially viewed a random-dot stereogram depicting a rectangle in front of a plane and coloured balls at different positions in depth. Next, they were presented with $+4.58'$ and $+13.73'$ faces and asked which one appeared to stick out more. Although these were not intended as formal tests of stereo vision, participants' ability to perceive them correctly indicated they all possessed some degree of stereo vision. These faces were either coloured or mid-grey, depending upon the condition.

Participants then completed two practice trials without feedback before beginning the task. These were randomly chosen examples from the experimental trials. The experiment consisted of 16 trials: 2 disparity magnitude of reference face ($|4.58'|$, $|13.73'|$) $\times$ 2 disparity sign of reference face ($+,-$) $\times$ 2 position of reference face (left, right) $\times$ 2 trial type (ascending, descending). Ascending and descending trials began with the comparison face set at a random point below or above the disparity magnitude of the reference face, respectively. For example, in an ascending trial with a $-13.73'$ reference face adjustments would begin with the disparity of the comparison face less than $+13.73'$. The order of within-subjects conditions was independently randomised for all participants. All trials were begun manually by the experimenter. Before participants began their adjustments the control monitor was switched off, so that the only light sources were the two display monitors.

Participants were instructed to adjust the depth of the comparison face until it appeared to 'stick out' the same amount as the reference face; they did this by pressing up and down arrow keys on a keyboard to increase and decrease its disparity-defined depth. When participants reported that the faces appeared to stick out the same amount, the PSE, quantified as the maximum angular disparity of the comparison face they chose as a match, was recorded manually by the experimenter. Participants were also instructed to inform the experimenter if they noticed anything unusual about the stimuli. This was to determine if the participants ever perceived the hollow faces as concave. They were then debriefed as to the true nature and aims of the study.

2.1.5 *Design.* The design was a 2 disparity magnitude of reference face ($|4.58'|$, $|13.73'|$) $\times$ 2 disparity sign of reference face ($+, -$) $\times$ 2 trial type (ascending, descending) $\times$ 2 reflectance (realistic colour, mid-grey) mixed design. All factors, except reflectance, were within-subjects. Position of reference face (left, right) was also included as a within-subjects factor in an initial analysis. The dependent variable was participants' PSE quantified as the maximum angular disparity of the comparison face chosen as a match. In trials where participants attempted to adjust beyond the maximum ($+24.78'$) or minimum ($0'$) comparison-face disparities, the trial was assigned the corresponding maximum/minimum value.

2.2 *Results and discussion*

In 31 out of the total of 352 trials conducted, participants attempted to adjust beyond the maximum (5 trials) or minimum (26 trials) disparity comparison faces. The assumption of normality was violated. While ANOVA was assumed to be robust, significant effects were also tested non-parametrically. All z values are corrected for ties. Initial analyses showed no effect of left/right position and data were collapsed across this factor. The ANOVA showed a significant disparity sign by magnitude interaction ($F_{1,20} = 16.88$, $p < 0.001$)—see figure 2, and magnitude by trial type interaction ($F_{1,20} = 9.54$, $p = 0.006$).

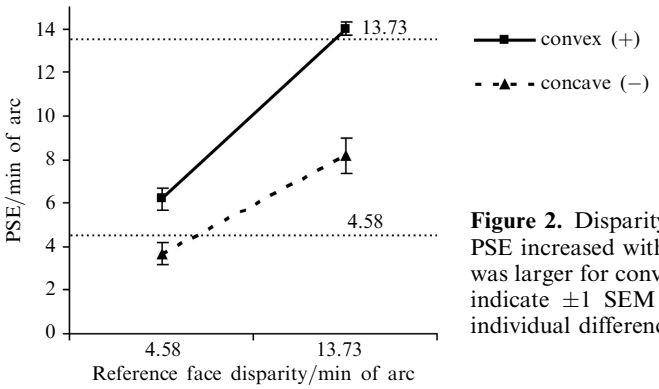

Figure 2. Disparity sign by magnitude interaction. PSE increased with disparity magnitude. This effect was larger for convex than concave faces. Error bars indicate ±1 SEM (corrected to remove effects of individual differences).

Analyses of simple main effects and corresponding Wilcoxon signed rank tests (see table 1 for descriptive statistics) showed that PSE estimates were higher for the |13.73′| reference faces than for the |4.58′| reference face in both positive ($F_{1,21} = 154.23$, $p < 0.001$; $z = 4.11$, $p < 0.001$) and negative ($F_{1,21} = 87.69$, $p < 0.001$; $z = 4.08$, $p < 0.001$) disparity sign conditions. The effect of magnitude for the hollow (negative sign) faces was consistent with the 'sign veto' hypothesis with exaggerated disparity faces appearing deeper. Inspection of figure 2 suggests that the interaction was due to the effect of disparity magnitude being greater when sign was positive than when it was negative.

Table 1. Cell means (M), standard deviations (SD), and medians (Mdn) for simple main effects analysis of the sign by magnitude interaction. All units are min of arc.

	Magnitude |4.58′|			Magnitude |13.73′|			Total		
	M	SD	Mdn	M	SD	Mdn	M	SD	Mdn
Convex (+)	6.31	2.34	5.51	14.04	1.60	14.22	10.09	1.33	9.68
Concave (−)	3.67	2.43	2.53	8.21	3.94	7.57	5.97	3.07	5.37
Total	4.95	2.02	4.17	10.71	2.20	11.29			

One-sample t-tests and sign tests showed that for +13.73′ reference face the matched comparison face was not significantly different ($t_{21} = 0.76$, $p = 0.456$; $n + 14$, $n - 8$, $p = 0.286$) while for the +4.58′ reference face the comparison was slightly overestimated ($t_{21} = 3.27$, $p = 0.004$; $n + 19$, $n - 3$, $p < 0.001$). This may in part be due to the inaccuracies in depth judgments observed in naive participants in experiments employing a stereoscope (McKee and Taylor 2010). PSE was higher for convex reference faces than for hollow reference faces irrespective of whether disparity magnitude was |4.58′| ($F_{1,21} = 23.14$, $p < 0.001$) or |13.73′| ($F_{1,21} = 44.47$, $p < 0.001$; $z = 4.11$, $p < 0.001$). This indicates that hollow faces tend to appear flatter than their convex counterparts consistent with the 'compression' effect observed previously (Hartung et al 2005; Kroliczak et al 2006). Flattening is a problem for the 'sign veto' hypothesis unless one posits that inconsistent sign reduces perceived depth but that, unlike the 'trade-off' hypothesis, this is independent of the associated disparity magnitude.

There were simple main effects of disparity magnitude, confirmed by Wilcoxon signed rank comparisons, for both ascending ($F_{1,21} = 149.94$, $p < 0.001$; $z = 4.11$, $p < 0.001$) and descending ($F_{1,21} = 137.75$, $p < 0.001$; $z = 4.11$, $p < 0.001$) trials. The effect of direction, with PSE higher in descending than ascending trials, is typical of the method of adjustment. Inspection of figure 3 suggests the magnitude by trial-type interaction was due to the effect of trial type being greater when disparity magnitude is exaggerated.

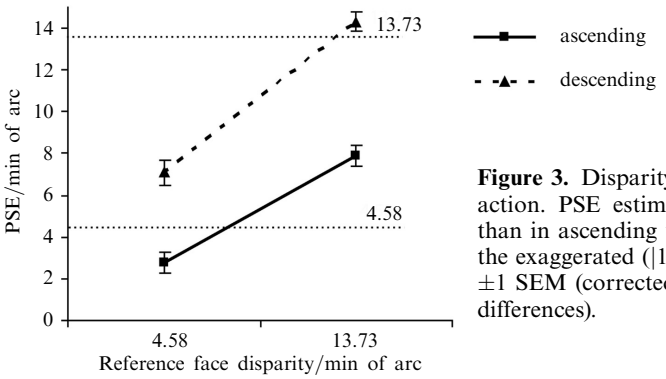

Figure 3. Disparity magnitude by trial-type interaction. PSE estimates were higher in descending than in ascending trials. This effect was greater for the exaggerated (|13.73′|) faces. Error bars indicate ±1 SEM (corrected to remove effects of individual differences).

There was no main effect of reflectance or any interactions involving this factor. This may suggest that prior knowledge does not concern typical depth magnitude of faces or 3-D structure as has been suggested previously (Hill and Johnston 2007), but, rather, simply determines perceived depth order. If prior knowledge does not concern depth magnitude, there would be no reason to expect that increasing its influence (through adding realistic colour to the face) would affect perceived depth magnitude. The null effect may also be due to a lack of sensitivity.

One participant reported after the experiment that the percept of the hollow face 'flipped' between convex and concave, but stated that the adjustments had been based upon the convex percept. Five participants reported that sometimes the faces appeared to be at different distances. As this was always reported afterwards, despite instruction to report during the experiment, it is difficult to determine what factors may have caused this. As participants were instructed to make the faces 'stick out' by the same amount, this distance effect may have introduced a confound, as the instruction could easily be misinterpreted as referring to the apparent proximity of the faces.

In experiment 1 we observed that as disparity magnitude associated with the hollow face increased so did perceived depth of the illusory face. This is concordant with the 'sign veto' hypothesis; namely that information from disparity magnitude is used to determine perceived depth magnitude of the illusory face even when depth-order information does not contribute to perceived convexity/concavity. There was no evidence of an effect of surface colour on perceived depth magnitude of the illusion.

3 Experiment 2

The primary aim of experiment 2 was to replicate the effect of disparity magnitude on the perceived depth of the illusory convex face by using the method of constant stimuli. We also used figural inversion as a manipulation of prior knowledge.

3.1 *Method*

3.1.1 *Apparatus and stimuli.* The Wheatstone setup was used as described for experiment 1 except that an Apple Macintosh was used instead of a PC so as to enable presentation and control of stimuli with MATLAB and PsychToolbox (Brainard 1997; Pelli 1997). Stereo pairs of images were generated in the same way as for experiment 1 except that maximum disparity of comparison faces ranged from 0′ to +18.36′ in increments of 1.15′. The reference faces used were the two coloured hollow faces from experiment 1: maximum disparities −4.58′ and −13.73′. Convex reference faces were omitted to limit trial numbers, but both convex–convex and concave–concave catch trials were employed showing one face with maximum disparity ±4.58′ and one with ±13.73′. All stimuli had realistic colouring and inverted stimuli were produced by flipping the image about the horizontal axis.

3.1.2 *Participants.* Participants were three male and fourteen female undergraduate students from the University of Wollongong aged between 18 and 55 years ($M = 21.70$ years, SD = 8.87 years). All reported 20/40 or better vision and received course credit for participating.

3.1.3 *Procedure.* Before beginning, the demands of the task were explained and written consent obtained. As in experiment 1, participants were not informed that any of the faces would be hollow. They were instructed that occasionally the faces may appear as though they were at different distances from them, but to make their judgments on the basis of the apparent depth of the face itself, described as the distance between the tip of the nose and the back of the head. This change in instruction was intended to avoid any ambiguity between depth and the apparent distance of the faces associated with the words "sticking out". Participants were seated and calibration of the stereoscope was performed as before. Images were displayed and responses recorded with MATLAB and Psychophysics Toolbox (Brainard 1997). They were positioned and scaled so as to match the stimuli in experiment 1. For each trial participants indicated which face appeared deeper, the left or right, responding by pressing '1' or '2' on the keyboard, respectively. Before commencing the experiment, participants completed 12 practice trials, randomly selected from the experimental trials. After this, all extraneous light sources were switched off and the experimenter left the room after activating the MATLAB experimental program. The experiment consisted of 408 experimental trials: 2 disparity magnitude of reference face ($-13.73'$, $-4.58'$) $\times$ 17 disparity magnitude of comparison face ($0'$ to $+18.35'$ in increments of $1.15'$) $\times$ 2 orientation (upright, inverted) $\times$ 6 repetitions. Trials were split into upright and inverted blocks. All faces in a block were shown in the same orientation. Both block and trial order were fully randomised for each participant independently. Within each set of 34 experimental trials one convex–convex and one concave–concave catch trial was included in which participants judged between the exaggerated and unexaggerated face. Although, again, no formal tests of stereo vision were conducted, performance on the convex–convex catch trials (all performed at or better than chance) indicated participants could distinguish between depths implied by binocular disparities. After the experiment, participants were asked if they had noticed anything unusual and then debriefed as to the true aims and nature of the study.

3.1.4 *Design.* The design was a 2 disparity magnitude of reference face ($-4.58'$, $-13.73'$) $\times$ 17 disparity magnitude of comparison face ($0'$ to $+18.35'$ in increments of $1.15'$) $\times$ 2 orientation (upright, inverted) repeated-measures design. The dependent variable was the percentage of trials when concave reference face was judged deeper than the convex comparison face.

3.2 *Results and discussion*

The assumption of normality was violated and non-parametric tests were used in addition to an ANOVA. Where sphericity was also violated Greenhouse–Geisser adjustments are reported. A preliminary analysis showed that block order had no effect and data were collapsed across this factor.

A repeated-measures ANOVA revealed a significant interaction between orientation and the disparity magnitude of the comparison face ($F_{6.49,\,173.82} = 2.23$, $p = 0.041$) together with a significant main effect of reference-face magnitude ($F_{1,\,16} = 7.80$, $p = 0.013$).

As expected, the percentage of "deeper" responses to the hollow reference face decreased as the disparity magnitude of the convex comparison face increased. However, even compared to comparison faces with 0 disparity, the hollow reference faces were only judged to be "deeper" 44.63% of the time. The effect of comparison-face disparity interacted with orientation but was significant for both upright ($F_{4.64,\,74.19} = 8.42$, $p < 0.001$)

and inverted faces ($F_{4.51,72.25} = 3.74$, $p = 0.006$). Overall, upright hollow faces were judged as deeper ($M = 29.29\%$, SD $= 20.52\%$; Mdn $= 33.82\%$) more than figurally inverted hollow faces ($M = 25.66\%$, SD $= 21.14\%$; Mdn $= 28.92\%$), Wilcoxon signed rank comparison $z = 2.73$, $p = 0.006$. This shows that figural orientation can affect perceived depth. Inspection of figure 4 suggests that the effect of figural orientation dissipates when comparison face disparity is high and 20% "deeper" responses may represent a lower asymptote. This may reflect a relatively high lapse rate due to the uncertainty associated with making judgments of relative depth.

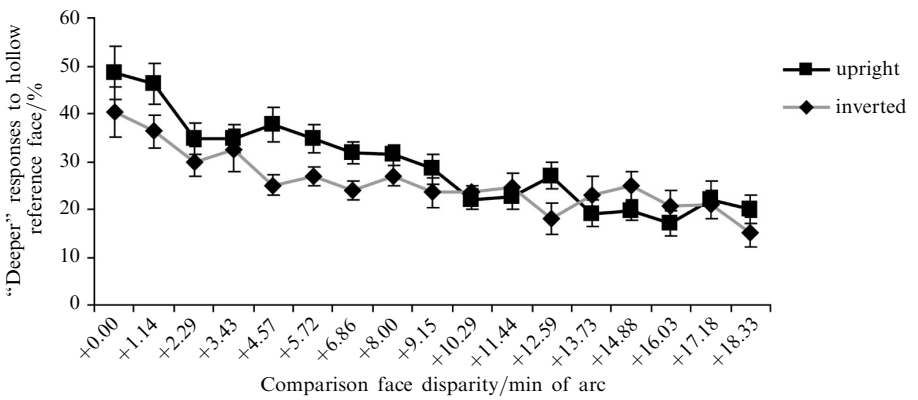

Figure 4. Orientation by magnitude of comparison-face interaction. Data are collapsed across magnitude of the hollow reference face. Upright hollow faces were judged as deeper more frequently than inverted hollow faces, although this effect is attenuated at larger comparison-face disparity magnitudes. Error bars indicate ± 1 SEM (corrected to remove effects of individual differences).

"Deeper" responses were significantly more frequent for the $-13.73'$ reference face ($M = 30.10\%$, SD $= 21.67\%$; Mdn $= 37.75\%$) than the $-4.58'$ reference face ($M = 24.85\%$, SD $= 20.47\%$; Mdn $= 20.59\%$) as shown by both the ANOVA and Wilcoxon signed rank comparisons ($z = 2.44$, $p = 0.015$). This is consistent with the results of experiment 1 in indicating that the hollow face with larger disparity magnitude looks deeper than the hollow face with smaller disparities, as predicted by the 'sign veto' hypothesis.

There was also a marginally significant interaction between magnitude of the reference face and orientation ($F_{1,16} = 3.58$, $p = 0.077$). Inspection of the relevant means showed that the effect of orientation was greater for the $-13.73'$ than the $-4.58'$ reference face (mean difference 5.65% and 1.61%, respectively). This suggests that the disparity of the hollow face is more likely to affect perceived depth when the face is upright, although "deeper" responses for $-4.58'$ may also be close to floor.

Collectively, participants performed better than chance, as shown by one-sample t-tests and sign tests, for convex ($M = 80.88\%$, SD $= 19.73\%$; Mdn $= 83.33\%$)[3] ($t_{16} = 7.55$, $p < 0.001$; $n + 15$, $n - 2$, $p = 0.002$), but not concave ($M = 51.66\%$, SD $= 29.88\%$; Mdn $= 58.33\%$) ($t_{16} = 0.75$, $p = 0.751$; $n + 7$, $n - 10$, $p = 0.629$) catch trials. This is consistent with a relatively high error rate even for convex faces where disparities differ by a factor of three. The level of uncertainty associated with the decision is even greater for concave–concave comparisons. Upon completion of the experiment, four participants reported that occasionally the faces appeared as though they were at different distances. None reported the faces ever appearing hollow.

Experiment 2 provided further evidence that disparity magnitude contributes to perceived depth magnitude even when disparity sign does not determine perceived

[3] Descriptive statistics are percentage 'correct'. For both concave and convex catch trials the 'correct' response was defined as regrading the exaggerated disparity face as deeper.

convexity/concavity. There was also evidence that weakening the influence of prior knowledge by figural inversion of the hollow face reduces the perceived depth of the illusory convex face.

4 General discussion

The hollow-face illusion persists even when the available binocular disparity information indicates unambiguously that the surface is actually concave. The primary aim of this study was to determine what happens to this information. Both experiments 1 and 2 showed that hollow faces with exaggerated disparities generate illusory faces that appears to have greater depth than those with unexaggerated disparities. This was measured by comparing hollow-face stimuli to a variety of convex faces with more/less disparity-defined depth. Our results were most consistent with the 'sign veto' hypothesis outlined in the introduction, which suggests that disparity magnitude can contribute to the perceived depth magnitude of the face even when the face percept is not consistent with the depth order as indicated by disparity sign. While adding realistic colour did not have any significant effect on perceived depth magnitude, upright hollow faces were more likely to be judged as 'deeper'. The latter finding was consistent with prior object knowledge affecting perceived depth of the illusory convex face.

The suggestion that disparity magnitude can affect perceived depth magnitude despite the perceived depth order being inconsistent with disparity sign is largely unprecedented. However, a similar effect has been observed in experiments with aniseikonic lenses used to manipulate binocular disparity in such a way as to imply slant in surfaces that are actually frontoparallel to the observer (Ogle 1950). Gillam (1967) and Seagrim (1967) have both observed that often, although the viewed surface is perceived as slanted, as disparity magnitude would predict, the direction of perceived slant is frequently the opposite to that implied by disparity sign. However, as neither study quantified the magnitude of perceived slant, it cannot be determined whether it was a direct function of disparity magnitude. In other work, where participants judged the depth order and depth magnitude of LEDs in a dark railway tunnel (Palmisano et al 2010), depth magnitude estimates appeared to scale with disparity magnitude even in instances when the depth order was misperceived. (These findings were auxiliary to the study and were not reported.)

"Mental pseudoscopy" (Yellott 1981; Yellott and Kaiwi 1979) as outlined in the introduction could provide an explanation for these phenomena and for our results. It is possible that the sign of binocular disparity is reversed mentally via reverse interpretations of left and right images. As it has been presented previously, the mental pseudoscopy explanation assumes that this reversal is applied to the entirety of the left and right images, predicting apparent depth reversal of the entirety of the viewed scene. Compelling evidence against this formulation consists of the ability to perceive illusory depth order reversals for some objects (eg hollow faces) whilst not others (eg stereograms, apparent direction of motion of the Pulfrich pendulum) at the same time in the same scene (Yellott 1981; Yellott and Kaiwi 1979). However, we suggest that binocular disparity may be interpreted in an object-based rather than retinal-based coordinate system. This would enable depth order information from sign to be interpreted (and in some cases reversed) for individual objects independently of the scene. This possibility could be characterised as "object-level mental pseudoscopy".

It is possible that, while relative disparities provide information about relative depth magnitude between object features, prior knowledge of faces as convex informs interpretation of their depth-order relationships. This may be similar to the way unmatched binocular features characteristic of da Vinci stereopsis can influence the perception of depth order in stimuli where disparity information is ambiguous (Anderson and Nakayama 1994).

Regarding the manipulations of object knowledge, the results are less clear. Adding surface colour information in experiment 1 had no significant effect on perceived depth. This is consistent with prior knowledge only affecting perceived depth order and not magnitude. It is also possible the high contrast edge information associated with surface colour may have improved the perception of depth magnitude from stereopsis (Bülthoff and Mallot 1988) offsetting any effect of prior knowledge (Hill and Johnston 2007).

If prior knowledge biases perception towards a face with typical rather than exaggerated depth, we might have expected the exaggerated depth faces to have appeared less exaggerated than they actually are when both convex and concave. However, as our experiment involved matching to convex comparison faces, this would not have been observable for convex reference faces, as both reference and comparison faces would appear equally flattened. Thus, this would have only been observable for hollow faces. The disparity sign by disparity magnitude interaction shown in figure 2 suggests that $13.73'$ reference faces were more flattened compared to their convex counterparts than $4.58'$ faces but this was not dependent on the presence of surface colour.

In experiment 2 there was a significant effect of figural orientation, with upright faces appearing deeper than inverted faces, at least when comparison faces had relatively low disparities. While this does suggest an effect of object knowledge on perceived depth magnitude, the effect did not interact significantly with the disparity magnitude of the reference face and so does not provide clear evidence in support of a bias towards any particular, presumably typical, depth. Determining the role of prior knowledge clearly requires further research, and figural inversion provides the more promising and clear manipulation for this. Comparison with a non-face depth probe would also ensure that the influence of prior knowledge does not also affect the scale on which perceived depth is quantified.

In the experiments reported disparity magnitude was manipulated by varying the intercamera distance used. This raises the possibility that the magnitude of monocular differences in rotation rather than disparities underlies the differences found between $-4.58'$ and $-13.73'$ reference faces (these monocular differences are visible in figure 1). This is unlikely as, in experiment 1, hollow faces were perceived as flatter than convex faces of the same disparity magnitude, consistent with previous findings. This flattening cannot be accounted for in terms of monocular rotation, as the only difference between convex and concave faces shown here was left/right reversal of the same images. It was also reported by participants in both experiments that sometimes the faces appeared to be at different egocentric distances. It is possible that this also co-varied with disparity although we did not investigate this. However, this is unlikely to account for the entirety of the disparity magnitude effect because (i) different apparent distances were only reported by a minority of participants, and (ii) in experiment 2 participants were explicitly instructed to ignore apparent distance and instructions were neutral with regard to the apparent proximity of the faces.

It was commented by one reviewer that perceptions of the hollow face often alternate between illusory convex and veridical concave. In fact, the relative predominance of either percept has been used as a method for quantifying illusion strength (Papathomas and Bono 2004). Participants in our study were asked to report anything unusual, and only one out of a total 39 participants reported ever experiencing the faces as hollow (none of the participants was informed of this possibility prior to the experiment). However, we cannot be sure that participants always made their judgments while perceiving the hollow face as convex (ie brief periods when they were not experiencing the illusion may have influenced their responding). In a brief follow-up experiment eight participants (including six undergraduates and two of the authors) were shown trials for 1 repetition of the upright condition from experiment 2, but with instructions

to indicate if the left, right, or neither of the faces appeared hollow for any or all of the time. They reported that in 34.93% of trials the hollow face was perceived as concave for at least part of the time. This was reported in 26.47% of trials involving the unexaggerated-disparity hollow face and 43.38% of trials involving the exaggerated-disparity face (this difference was significant—Wilcoxon signed rank $z = 2.03$, $p = 0.042$). The difference between conditions suggests that increasing disparity magnitude helps disambiguate the illusion, in difference to previous findings (Hill and Johnston 2007).

If participants often perceived the reference face as hollow, this might present a problem when interpreting the results of experiment 2. Here participants indicated the "deeper" face and, since "deeper" can equally well refer to concave or convex depth, participants could have legitimately compared the convex depth of the comparison face to the concave depth of the reference face. However, given our own experience and the lack of any participant reports to that effect, we feel that this is very unlikely. Even though our participants could, when explicitly asked, identify the presence of reversed disparities 35% of the time, it seems unlikely that anyone consistently experiencing one of the faces as hollow, and responding on the basis of that percept, would not have reported it. It is likely that this 35% detection rate represents the worst-case scenario, since it was obtained under conditions where the experimental demands would have biased participants to see a hollow face (as this was the only type of response they were asked to make). In summary, this study aimed to determine how, if at all, information from binocular disparity is used when viewing the hollow-face illusion. We present evidence that disparity-magnitude information affects perceived depth magnitude even when depth order information from disparity sign is disregarded.

Acknowledgment. This research was supported by Australian Research Council Grant DP0986898.

References

Anderson B, Nakayama K, 1994 "Toward a general theory of stereopsis: Binocular matching, occluding contours and fusion" *Psychological Review* **101** 414 – 445

Brainard D, 1997 "The psychophysics toolbox" *Spatial Vision* **10** 433 – 436

Bülthoff H, Mallot H, 1988 "Integration of depth modules: stereo and shading" *Journal of the Optical Society of America* **5** 1749 – 1757

Georgeson M, 1979 "Random-dot stereograms of real objects: observations on stereo faces and moulds" *Perception* **8** 585 – 588

Gillam B, 1967 "Changes in the direction of induced aniseikonic slant as a function of distance" *Vision Research* **7** 777 – 783

Gregory R, 1973 "The confounded eye", in *Illusion in Nature and Art* Eds R L Gregory, E H Gombrich (London: Duckworth) pp 49 – 96

Gregory R, 1980 "Perceptions as hypotheses" *Philosophical Transactions of the Royal Society of London, Series B* **290** 181 – 197

Hartung B, Schrater P R, Bülthoff H H, Kersten D, Franz V H, 2005 "Is prior knowledge of object geometry used in visually guided reaching?" *Journal of Vision* **5** 504 – 514

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, Bruce V, 1996 "Effects of lighting on the perception of facial surfaces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 986 – 1004

Hill H, Johnston A, 2007 "The hollow-face illusion: Object-specific knowledge, general assumptions or properties of the stimulus?" *Perception* **36** 199 – 223

Kroliczak G, Heard P, Goodale M, Gregory R, 2006 "Dissociation of perception and action unmasked by the hollow-face illusion" *Brain Research* **1080** 9 – 16

McKee S, Taylor D, 2010 "The precision of monocular and binocular depth judgements in natural setting" *Journal of Vision* **10** 1 – 13

Norman J, Todd J, 1995 "The perception of 3D structure from contradictory optical patterns" *Perception & Psychophysics* **57** 826 – 834

Ogle K, 1950 *Researches in Binocular Vision* (New York: Hafner)

Palmisano S, Gillam B, Govan D, Allison R, Harris J, 2010 "Stereoscopic perception of depths at large distances" *Journal of Vision* **10** 1 – 16

Papathomas T, Bono L, 2004 "Experiments with a hollow mask and a reverspective: Top – down influences in the inversion effect for 3-D stimuli" *Perception* **33** 1129 – 1138

Pelli D, 1997 "The VideoToolbox software for visual psychophysics: Transforming numbers into movies" *Spatial Vision* **10** 433 – 436

Seagrim G, 1967 "Stereoscopic vision and aniseikonic lenses" *British Journal of Psychiatry* **58** 337 – 350

Walker J, 1986 "The hyperscope and the pseudoscope aid experiments on three-dimensional vision" *Scientific American* **255** 124 – 130

Wheatstone C, 1852 "Contributions to the physiology of vision—Part the second. On some remarkable and hitherto unobserved phenomena of binocular vision" *Philosophical Transactions of the Royal Society of London* **142** 1 – 17

Yellott J I Jr, 1981 "Binocular depth inversion" *Scientific American* **245**(1) 118 – 125

Yellott J I Jr, Kaiwi J I, 1979 "Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse perspective" *Perception* **8** 135 – 142

© 2011 a Pion publication

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

PERCEPTION

VOLUME 40 2011

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.