Perception, 2007, volume 36, pages 461 – 470

DOI:10.1068/p5549

Eye vergence is susceptible to the hollow-face illusion

Joachim Hoffmann, Albrecht Sebald
Department of Psychology, University of Würzburg, Röntgenring 11, D 97070 Würzburg, Germany;
e-mail: hoffmann@psychologie.uni-wuerzburg.de

Received 22 November 2005, in revised form 29 May 2006; published online 21 February 2007

Abstract. When an observer looks at a hollow mask of a face, a normal convex face is often perceived [the hollow-face illusion — Gregory 1973, in *Illusion in Nature and Art* (London: Duckworth) pp 49 – 96]. We show that in exploring an illusory face, the eyes converge at the illusory and not at the real distances of fixated targets like the tip of the nose. The 'vergence error' appears even though the resulting disparities of the two retinal images of the target provide feedback that would allow an immediate correction. It is presumably the success of recognising a familiar object (a face) which overrides the correction of convergence. This suggests that the brain strives for a congruency of eye vergence and distance perception.

1 Introduction

When an observer looks at the inside of a mask or mould of a face from a certain distance, the mask often appears as a normal convex hemispherical face (the hollow-face illusion — Gregory 1973). The strength of the illusion has been shown to be influenced by stereoscopic as well as by pictorial information: the illusion is rarely seen if the hollow mask is viewed from short distances, ie if retinal disparities unambiguously indicate its concave shape. Conversely, its appearance is facilitated under monocular viewing conditions, ie when no stereoscopic depth information is available (Hill and Bruce 1993). Furthermore, the more the pictorial information corresponds to the familiar pattern of a face, the more likely the illusion will occur. In particular, the illusion is stronger for upright than for inverted faces, especially if the (upright) illusory face appears to be illuminated from above (ie the hollow mask is illuminated from below), as faces usually are (Sakurai et al 1985; Hill and Bruce 1993). Finally, the illusion of seeing concave displays as convex is much weaker with moulds of less familiar forms like potatoes (Johnston et al 1992; Hill and Bruce 1994), and it disappears altogether if all monocular pictorial cues are eliminated (as in random-dot stereograms of face masks — cf Georgeson 1979).

The hollow mask provides various, partly ambiguous, sources of depth information including motion parallax, self occlusions, perspective, and shading. However, it is monocular pictorial cues that indicate the presence of a face, and binocular stereoscopic cues that probably have the strongest impact on what is perceived. Whereas the pictorial cues reactivate face representations that provide knowledge about convex shape, the stereoscopic cues indicate the concave shape of the mask. The way the contradiction is resolved depends on which information currently dominates perception: the illusion of a convex face emerges if the reactivated knowledge prevails, and a hollow mask is correctly perceived if the stereoscopic information prevails (Gregory 1973; Georgeson 1979).

Two recent studies have shown that this contradiction between reactivated knowledge and stereoscopic cues not only misguides perception but also affects pointing or reaching movements. Hartung et al (2005) presented computer-generated images of normal and hollow faces for three different tasks. In the 'verbal task', participants were asked to give a verbal estimate of the distance from their viewing position to either the nose or the cheek of the faces. In the 'non-haptic reaching task', participants were asked to 'touch' either the nose or the cheek of the (virtual) face. Shortly after movement onset, the image of the face was removed so that the reaching had to be carried out without visual feedback. Finally, a so-called 'haptic task' was almost identical to the 'non-haptic task', except that participants always received false haptic feedback at the tip of the index finger as if they actually had been touching the surface of a face. Accordingly, every reaching was fed back as being correct, irrespective of how far it extended.

Under these conditions it was found that performance in all three tasks was similarly affected by the hollow-face illusion: although looking at a hollow face participants estimated the cheeks as being further away than noses and they reached to cheeks further than to noses, regardless of whether or not haptic feedback was provided. In comparison to normal faces, though, the differences between noses and cheeks were reduced in all three tasks, indicating that the binocular information was not totally disregarded. The authors concluded that both distance estimates and hand movements rely on depth information, which results from a combination of current binocular cues and reactivated knowledge about the normal shape of faces, with the impact of knowledge dominating.

In the second study Króliczak et al (2006) presented either a convex or a concave (hollow) mask of a female face. The faces were mounted on a 'reference plate', such that the normal (convex) face protruded in front and the hollow (concave) face receded behind it. Additionally, a small cylindrical magnet was placed at the side of the cheek or at the forehead as the target. Participants had to accomplish three different tasks. In the 'paper-and-pencil drawing task', participants were to indicate the distance of the current target by placing a mark to the right (for near) or to the left (for distant) of a vertical reference line representing the reference plate of the mask. In the 'slow pointing task' participants had to point either directly to the location where they perceived the target or to the corresponding distance below the face. Finally, in the 'fast flicking task' participants were asked to flick a small magnet off from the face as quickly and accurately as they could, using their index finger. Both pointing and flicking were performed without visual feedback, because the view of the face was blocked by LCD (PLATO) goggles as soon as the movement started. However, note that haptic feedback was always available in the flicking task because participants, of course, noticed whether or not they had hit the target.

The data of the drawing task revealed a robust hollow-face illusion, ie, when looking at a hollow face, participants' estimates of target distances matched those they gave when looking at a normal face. Likewise, the final positions of the pointing movements made to the hollow face were clearly in front of the reference plate. However, as in the study of Hartung et al (2005), the differences between the estimated distances as well as between the final pointing positions for the cheek and forehead targets were compressed in comparison to the normal face, again indicating that binocular information still maintains some influence on perception and pointing.

In sharp contrast to the data from 'drawing' and 'pointing', the flicking movements were directed to the real positions of the targets. Thus, when looking at the hollow face participants saw and pointed to the targets at positions in front of the reference plate (the illusion) but their flicking movements were directed to positions behind the reference plate. In other words, whereas perception and pointing movements were misled by the hollow-face illusion, the flicking movements were not.

Altogether, both studies reveal that some movements are susceptible to the hollow-face illusion, whereas other movements remain immune to it. Two reasons have been discussed to account for this contradiction. On the one hand, Króliczak et al (2006) refer to the fact that the movement time for pointing was more than three times as long as the movement time for flicking (1660 ms versus 471 ms). Accordingly, they assume that flicking relied on automatic visuo-motor control mechanisms which are not influenced by top–down information like reactivated face knowledge. In contrast, the slow pointing movements, they assume, are mediated by control mechanisms which are subject to cognitive factors as, in the present case, the seen (illusory) distances of the targets.

On the other hand, it should be noted that only the flicking movements produced veridical feedback about success or failure. Króliczak et al (2006) reported that participants failed to hit the target on about half the trials when flicking targets from the hollow face. This high frequency of failures possibly caused binocular information to gain more impact on movement control so that the influence of reactivated face knowledge was repressed.

The present study was not designed to clarify this issue, but rather explores the impact of the hollow-face illusion on another face-oriented movement, the vergence movements of the eyes. Just as the preceding studies explored where the hands move if they reach to a target on an illusory face, in the present experiments we explore, to our knowledge for the first time, where the eyes gaze when observers look at an illusory face. More precisely, we ask at which distance the eyes converge if observers are instructed to fixate on the tip of the nose of a hollow mask: do the eyes diverge to the far location of the tip of the real nose or do they converge to the near location of the illusory nose (figure 1)?

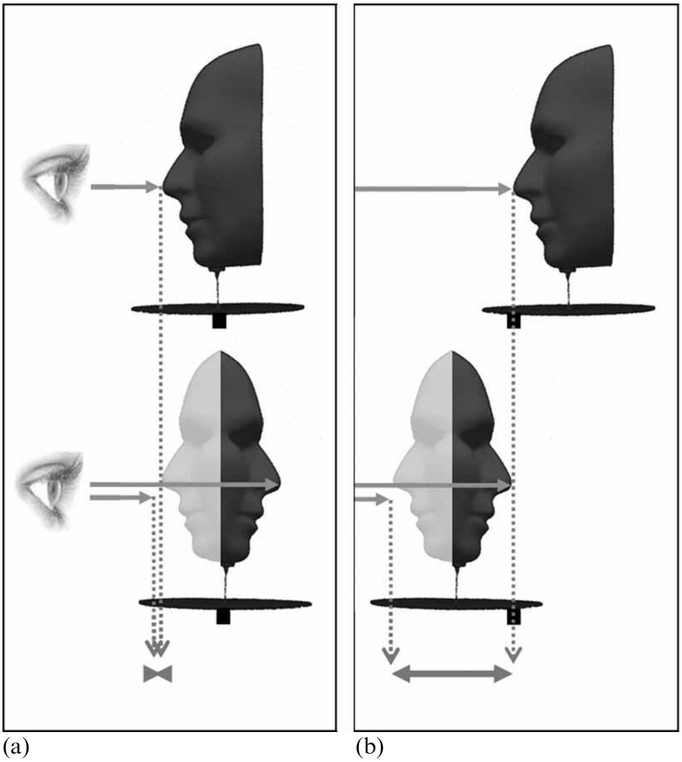
(a) (b)

Figure 1. Illustration of mask presentation. In experiment 1 the mask rotated around a vertical axis in the centre of the virtual head (a). In experiment 2 the mask rotated around a vertical axis going through the tip of the nose (b). The upper figures illustrate masks presented in front view to the observer (0°). The lower figures show masks presented in back view to the observer (180°) whereby the appearance of an illusory face is foreshadowed. In experiment 1, the tip of the nose of the illusory face is approximately equidistant to the one of the real faces in the front view. In experiment 2, the tip of the illusory nose appears at a closer distance to the observer than the tip of the real nose in the front view.

Vergence movements of the eyes fixating on a target are certainly highly automated responses. Furthermore, the visual system immediately provides veridical feedback whether the desired target is fixated or not. In the case of success, the target becomes centred in the fovea of both eyes, whereas in the case of failure, the retinal images of the target would be projected on disparate retinal locations. Both features, movement automation and the availability of immediate feedback, lead us to expect that vergence movements, like the flicking movements in the study by Króliczak et al (2006), will not be susceptible to the hollow-face illusion.

Moreover, there is evidence showing that vergence movements of the eyes are initiated by local disparities largely independent of depth perception (Mitchell 1970; Masson et al 1997). For example, they briefly presented his subjects with two target stimuli, one for each eye, with crossed and uncrossed disparities, so that either convergent or divergent eye movements were required. The stimuli always elicited vergence movements in accordance with the given disparities, even if the stimuli were too different to be fused, ie to create a stable percept. This direct initiation of the vergence movements by local disparities, irrespective of subjective perception, makes it very unlikely that vergence movements are susceptible to cognitive top – down influences.

In summary, the given evidence suggests that vergence movements will be immune to the hollow-face illusion. The following two experiments are aimed at an empirical confirmation of this supposition. However, the results will show that the vergence tends to adapt to illusory instead of to real locations.

2 Experiment 1

2.1 *Method*

2.1.1 *Convergence measurement.* Vergence movements of the eyes were recorded photoelectrically. Participants wore spectacles that were equipped with two diodes emitting pulsed infrared light and two infrared sensitive transistors for each eye. Diodes and transistors were adjusted at the nasal and temporal sides of each eye so that both transistors absorbed reflected light from the sclera and the iris. A rotation of the eye in the horizontal plane resulted in opposed changes of the amount of light absorbed by the two transistors. For example, in focusing on a close target, the less reflecting iris moves to the nasal side in both eyes such that there is less light to absorb by the nasal transistors but more light to absorb by the temporal transistors, ie the signal from the nasal transistors is weakened whereas the signal from the temporal transistors is strengthened. The signals were electronically processed in such a way that in each eye the signal from the nasal transistor was subtracted from the signal of the temporal transistor, and these two differences were added. Accordingly, this integrated convergence signal increased the closer the fixation point, and decreased with increasing distance of the fixated target.

In each individual trial the integrated convergence signal was recorded over 5 s with a sampling rate of 100 Hz beginning 1 s after a warning signal. In order to receive a measure of convergence as unbiased as possible, we decided to calculate the mean of the resulting 500 values (5 s) without any data cleansing (in particular because a tentative elimination of outliers did not change the results). After the 5 s, another tone signalled that recording was finished and that the gaze could be released. The adequacy of the used measurement was validated in a block of trials in which participants were to fixate on a cross at twelve different distances. Only those participants who showed a continuous decrease of the convergence measure with increasing distances of the cross were accepted in the final sample.

2.1.2 *Stimuli and procedure.* The experiments were performed in a small room of 7.30 m^2 (2.40 m by 3.05 m) that was illuminated by a single masked dim light of 40 W, mounted on the wall behind and somewhat above the head of the participants. Participants sat in a chair with the head placed on a chin-and-head rest in order to maintain a consistent viewing condition. Directly in front of the participants either a fixation cross or a monochrome face mask was presented at eye level against a black background. The fixation cross (3.2 cm high and 3.2 cm wide) was presented in black on a white paperboard (14.5 cm high and 21 cm wide). The paperboard with the cross was mounted on a sliding mechanism, which allowed us to present the cross at different distances. The face mask, which had approximately the colour of skin, was 23 cm long and 15.2 cm wide. The mask was also mounted on the sliding mechanism with a rod, which allowed us to rotate the mask. The distance between the axis of rotation and the tip of the nose was 3.9 cm.

Participants completed two blocks of trials with the order counterbalanced between them. In one block the fixation cross was presented once at each of twelve different distances from the observer, ranging from 20 to 100 cm in random order. In the other block the mask was presented with the axis of rotation at a distance of 50 cm from the observer. There was one presentation of the mask in each of the twelve different orientations, ranging from 0° (the straight view on the convex mask) over 180° (the straight view on the concave or hollow mask) to 330° in 30° increments. The order of presented orientations was randomised for each participant.

In each individual trial a warning signal (a tone) was presented at which time participants were asked to avoid blinking and to steady their fixation either on the cross in one block or on the tip of the nose of the mask in the other block of trials. After each presentation of the mask, participants were also asked to indicate whether or not they had clearly seen a normal convex face, so that trials in which the illusion appeared could be separated from trials in which the illusion did not appear. Finally, the fixation cross was adjusted at a new distance or the mask was adjusted at a new orientation for the next trial. The experiments lasted about 20 min.

2.1.3 *Participants.* In the first experiment sixty-eight students of the University of Würzburg took part as a fulfillment-of-a-course requirement. Eleven participants showed no systematic variation of the convergence measure with the distance of the cross and were discarded from further analyses.

2.2 *Results*

Figure 2 shows the results of the block in which the cross was to be fixated. The mean convergence measure of the fifty-seven remaining subjects decreases with increasing distance of the fixation cross, as expected.

The impact of the illusion on convergence can be assessed only under conditions that allow the illusion to appear. This applies to back views of the mask at 150°, 180°, and 210° (that is when participants look into a hollow mask). To reduce the impact of interindividual variability due to different fits of the spectacles, a difference measure was used (figure 1a). For each participant, the convergence measures at orientations of the mask of 150°, 180°, and 210° were subtracted from her/his convergence measure from fixating on the tip of the nose in the front view of the mask (0°, that is when participants look straight on a convex mask). In the case of no illusion (ie where participants correctly perceived the hollow side of the mask), a positive difference is to be expected because the tip of the nose is much closer in the front view than in any of the back views of the mask. In the case of an illusion (ie when participants perceive a convex face but the mask is hollow), the same result is to be expected provided that the convergence remains uninfluenced by the illusion; ie the eyes converge at the real distances of the tip of the nose despite the participants' perception of an illusory convex face. In contrast, if the eyes converge at the distance of the illusory nose, no differences between the convergence measures are to be expected because the illusory nose should be seen approximately at the same distance as the nose in the front view.

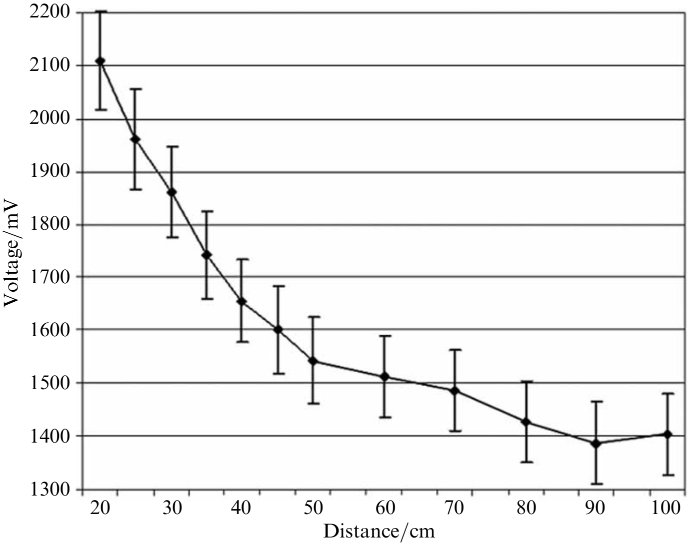

Figure 2. Mean convergence measures (±SE) from fixating on a cross at different distances. Eleven participants who showed no systematic co-variation of convergence with the distance of the cross were discarded from the sample.

Figure 3 shows the results: the white bars depict the means of the individual differences for all trials in which participants reported to have perceived a hollow mask instead of a convex face, ie trials in which no illusion was experienced (32, 28, and 31 cases at the orientations of the mask of 150°, 180°, and 210°, respectively). In all three orientations positive differences were recorded, indicating that in these trials, participants fixated at a more distant location when looking at the hollow mask compared to looking at the convex mask in the front view. For the mask orientations of 150° and 210°, the mean differences significantly deviated from zero [$t_{150}(31) = 2.427$; $t_{210}(30) = 2.899$; both ps < 0.05, one-tailed]. For the mask orientation of 180°, the mean difference fell short of the significance level [$t_{180}(27) = 1.517$, $p = 0.07$, one-tailed].

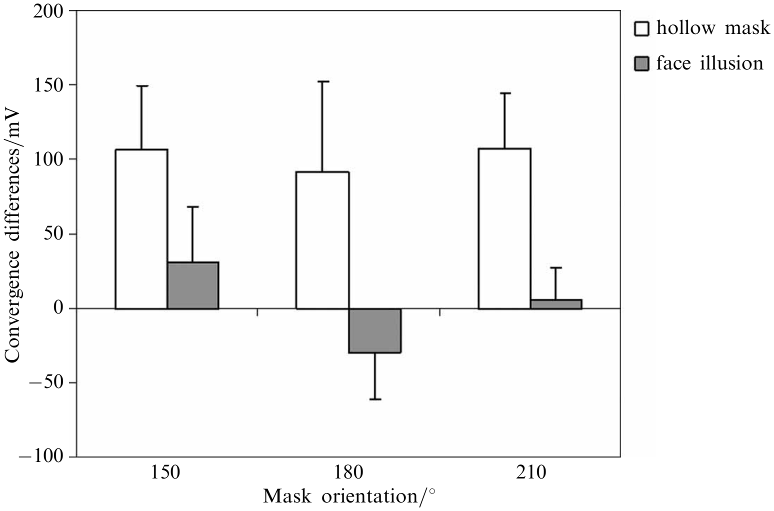

Figure 3. Mean differences (±SE) of the individual convergence measures recorded from the front view of the mask and from mask presentations at 150°, 180°, and 210°, separately for trials in which participants did (grey columns) and did not (white columns) experience a hollow-face illusion in experiment 1.

The grey bars depict the means of the differences for all trials in which participants reported having perceived an illusory convex face despite looking into a hollow mask (25, 29, and 26 cases at orientations of the mask of 150°, 180°, and 210°, respectively). In none of the three orientations did the difference significantly deviate from zero [$t_{150}(24) = 0.829$; $t_{180}(28) = -0.965$; $t_{210}(25) = 0.280$; all ps > 0.3, two-tailed). Thus, in the case of an illusion, participants converged at approximately the same distance as in fixating on the tip of the nose in the front view of the mask.

2.3 *Discussion*

In trials in which participants did not experience the hollow-face illusion (that is, when presented with a hollow mask, they correctly perceived that it was concave and not a convex face) the data revealed an appropriate adjustment of the vergence to the veridical distance of the tip of the nose. Participants who experience the illusion should produce the same results if the vergence of the eyes were not also affected by the illusion. This was not the case. In contrast, the data revealed that participants who reported to have seen a convex face despite looking at a concave mask fixated at approximately the distance of the tip of the nose in the front view of the mask, ie at the distance in which the tip of the illusory nose appeared to be. Thus, the illusion affected the vergence of the eyes in the same way as it affected perception. Experiment 2 was conducted in order to reassess this result under somewhat different conditions.

3 Experiment 2

3.1 *Method*

In experiment 2 the same conditions applied as in experiment 1 except that the mask rotated around a vertical axis which went through the tip of the nose instead of through the centre of the virtual head as in experiment 1 (figure 1b).

Accordingly, the distance of the tip of the nose to the observer remained constant irrespective of the orientation of the mask. In contrast to experiment 1, no difference between the convergence measures from the front and the back views of the mask now indicates that the convergence resists the illusion, whereas differences would indicate that the vergence movements are affected by the illusion. In particular, negative differences are to be expected because the illusory nose would now appear to be closer to the observer than the nose in the front view.

3.1.1 *Stimuli and procedure.* The same stimuli and procedure were applied as in experiment 1.

3.1.2 *Participants.* Seventy-five students of the University of Würzburg took part in experiment 2 as a fulfillment-of-a-course requirement. Again, eleven subjects showed no systematic increase of convergence with the distance of the cross and were discarded from further analyses.

3.2 *Results and discussion*

The data were processed in the same way as in experiment 1. Figure 4 shows the results of the block in which the cross was to be fixated. The mean vergence measures of the remaining sixty-four subjects decrease with the distance of the fixation cross, as expected.

In evaluating the data of the block in which the mask was presented, we again calculated the differences of the individual convergence measures between fixating on the tip of the nose from the front view of the mask (0°) and from its orientations of 150°, 180°, and 210°. In the case of no illusion, no differences are expected. In the case of an illusion, the same results are expected if the convergence remains uninfluenced by the illusion. However, if the eyes converge to the distance of the illusory nose, substantial negative differences are to be expected because the illusory nose should be seen at a closer distance than the nose from the front view.

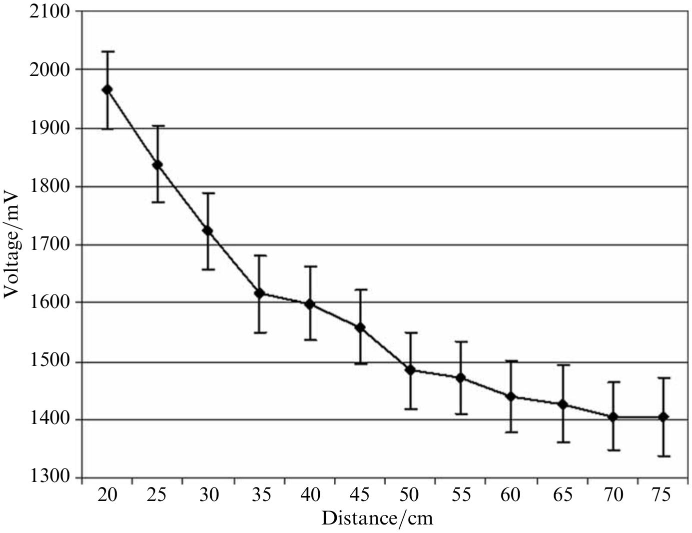

Figure 4. Mean convergence measures (±SE) from fixating on a cross at different distances. Eleven participants who showed no systematic co-variation of convergence with the distance of the cross were discarded from the sample.

Figure 5 shows the results. The white bars depict the means of the differences from all trials in which participants did not experience the illusion at the respective mask orientation (41, 27, and 43 cases at 150°, 180°, and 210°, respectively). The differences do not deviate from zero in any of the three orientations [$t_{150}(40) = 0.560$; $t_{180}(26) = 0.632$; $t_{210}(42) = 0.310$; all $ps > 0.5$, two-tailed)]. If participants did not experience the illusion, they always fixated at approximately the same distance regardless of the mask orientation. The grey bars depict the means of the differences from all trials in which participants reported that they had perceived a convex face (the illusion; 23, 37, and 21 cases at 150°, 180°, and 210°, respectively). In all three orientations negative

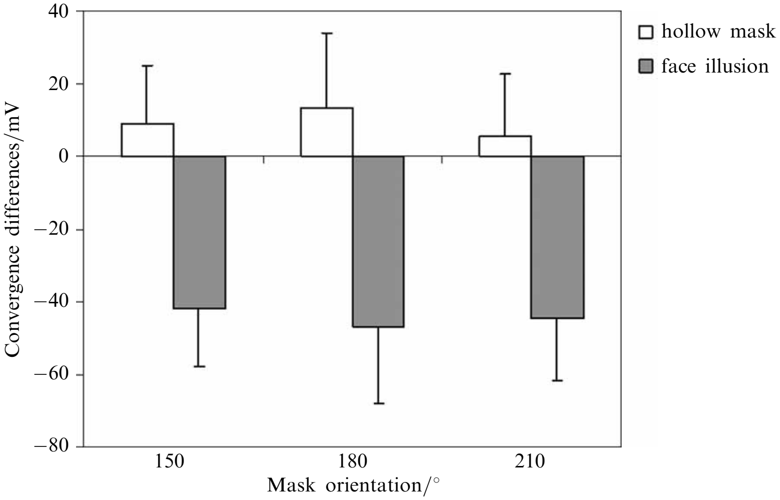

Figure 5. Mean differences (±SE) of the individual convergence measures recorded from the front view of the mask and from mask presentations at 150°, 180°, and 210°, separately for trials in which participants did (grey columns) and did not (white columns) experience a hollow-face illusion in experiment 2. In contrast to experiment 1 (figure 3), mean differences around zero now indicate that the convergence resists the illusion, whereas negative differences indicate that the convergence is susceptible to the illusion.

differences of the convergence measures resulted, which significantly deviated from zero [$t_{150}(22) = -1.988$, $t_{180}(36) = -2.269$; $t_{210}(20) = -1.854$; all ps < 0.05, one-tailed]. Thus, in the case of an illusion, participants converged to a closer distance than when they were fixating on the tip of the nose in the front view of the mask.

4 General discussion

Both experiments show that, under the present conditions, the vergence of the eyes and depth perception are not dissociated. Participants who perceived a convex face despite looking into a hollow mask fixated on the tip of the nose at its illusory distance. Participants who did not experience the illusion fixated on the tip of the nose at its real distance. In other words, the convergence is always adjusted so that the eyes converge at the distance where the target is subjectively located.

Two issues deserve further elaboration: first, our results seem to contradict the aforementioned control of vergence movements by local disparities irrespective of perception (Mitchell 1970; Masson et al 1997). The contradiction disappears, however, if one takes into account that only initial components of vergence movements have been shown to be directly determined by disparities. For example, Mitchell (1970, page 152) explicitly stated: "Eye movement recordings ... showed that on [longer] presentation of dissimilar targets there was typically an initial vergence eye movement in the correct direction [corresponding to the given disparities] which stopped well short of the amount required to binocularly fixate the two targets, after which the eyes would drift back to a position somewhere near the positions they adopted prior to the stimulus presentation". In the present experiments the onset of mask presentation was not clearly defined so that only late vergence eye movements could be recorded. Thus, it might well be that, in looking at a hollow mask, the vergence initially is veridically determined by binocular disparity, but becomes prompted by the illusory distance of the tip of the nose as soon as the reactivation of face knowledge gives rise to the appearance of an illusory convex face. Initial vergence movements to the real distances would be in full agreement with the recent finding of Króliczak et al (2006) that quick ballistic hand movements (flicking movements with latencies of about 500 ms) were veridically determined even though participants experienced the hollow-face illusion. One of the cues that Króliczak et al (2006) assume to drive the accurate flicking movements was certainly vergence which, according to our results, would have to be early vergence components. In order to explore this supposition, further experiments are needed in which the time course of vergence movements is recorded beginning with the onset of a hollow face.

Regardless of the initial vergence movements, a second issue concerns the fact that the eyes later adapt to a focus which is consistent with the perceptual illusion, even though the resulting disparities unambiguously indicate that the vergence is falsely adjusted. Typically, actions that fail to be successful are immediately recalibrated so that the intended goals are reached. For example, participants who wear distorting spectacles correct their behaviour appropriately even before the distortions are compensated for in perception (Kohler 1964), and eye movements have been shown to quickly adapt to target displacements even when the displacements are not consciously detected (Hajos and Fey 1982; Deubel 1987). In the present case, however, the falsely adjusted convergence has not been corrected. Instead, participants selected a fixation point well in front of the concave mask, possibly to make it easier to maintain the illusion by defocusing the image. In other words, the visual system tends to tolerate the disparities between the two retinal images in favour of maintaining a more recognisable image—an ordinary face. Thus, our data suggest that the success of recognition tends to override the failure of fixation.

Acknowledgments. We thank Georg Schüssler for designing and producing the spectacles for the measurement of eye convergence and three anonymous reviewers for valuable suggestions on an earlier version of the article.

References

Deubel H, 1987 "Adaptivity of gain and directions in oblique saccades", in *Eye Movements: From Physiology to Cognition* Eds J K O'Regan, A Levy-Schoen (Amsterdam: North Holland) pp 181 – 190

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

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

Hajos A, Fey D A, 1982 "Lernprozesse des okulomotorischen Systems" *Psychologische Beiträge* **24** 135 – 158

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

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365 – 375

Kohler I, 1964 "The formation and transformation of the perceptual world" *Psychological Issues* **3**(4) 1 – 173

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

Masson G S, Busettini C, Miles F A, 1997 "Vergence eye movements in response to binocular disparity without depth perception" *Nature* **389** 283 – 286

Mitchell D E, 1970 "Properties of stimuli eliciting vergence eye movements and stereopsis" *Vision Research* **10** 145 – 162

Sakurai K, Endo M, Maruyama K, 1985 "Depth reversal in the aperture viewing of facial shading patterns" *Tohoku Psychologica Folia* **44** 70 – 80

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

PERCEPTION

VOLUME 36 2007

www.perceptionweb.com