Failure of facial configural cues to alter metric stereoscopic depth

Barbara J. Gillam   School of Psychology, University of New South Wales, Sydney, NSW, Australia

Barton L. Anderson   School of Psychology, University of New South Wales and University of Sydney, Sydney, NSW, Australia

Farhan Rizwi   School of Psychology, University of New South Wales, Sydney, NSW, Australia

J. Burge, M. A. Peterson, and S. E. Palmer (2005) reported that an ordinal cue to depth can influence the perception of metric depth in stereoscopic displays. They argued that when a familiar figure—a face—is placed stereoscopically closer than a background there is greater perceived depth relative to the ground than when the face shape is placed stereoscopically further and becomes the ground. This result suggests the possibility that a non-metric depth cue—the familiarity of a figure—can influence the perception of metric depth in stereoscopic displays. However, the method leaves open the possibility that these results were due to a response bias, rather than from a genuine change in perceived depth. To assess this possibility, we used the same basic stimulus but directly measured the perceived depth difference between the face and non-face surfaces when arranged as figure and ground or ground and figure respectively using a separate double depth probe to measure perceived depth. We found no difference between the perceived depth of familiar and unfamiliar figures as a function of whether they were stereoscopically figure or ground. We conclude that the J. Burge et al. (2005) result depends on their particular task and is likely to reflect a response bias. It is premature to conclude that facial configural cues distort perception of metric depth although we argue that there are circumstances in which ordinal cues do influence metric depth.

Keywords: depth perception, figure-ground, disparity, configural cues, cue-combination, occlusion, familiarity

Introduction

Burge, Peterson, and Palmer (2005) reported two experiments designed to test whether ordinal depth information influences the extent of metric depth seen between two regions separated by an unambiguous binocular disparity. The ordinal information in their displays was specified by configural figure-ground cues, namely, the familiar outline of a face in profile. This has been found to bias figure-ground organization in 2-D displays so that the face tends to be seen as figure (Peterson & Gibson, 1994). The authors argue that any evidence that ordinal cues influence metric depth poses a problem for the "weak fusion" theory of cue combination proposed by Landy, Maloney, Johnston, and Young (1995) which assumes that different cues must be convertible to the same units for combination.

The figures used by Burge et al. (2005) are shown in Figure 1.

By varying the disparity, the side of the image containing the face profile can be placed either in front as figure or behind as ground. In the latter case the figure is a

meaningless shape because the convexities of the face become concavities on the non-face and vice versa. Burge et al. (2005) hypothesized that more depth would be seen when the face is placed stereoscopically in front (a cue consistent condition) than when it is placed stereoscopically behind (a cue inconsistent condition). Consistent and inconsistent stimuli were presented as sequential pairs. One member of the pair was presented at a fixed disparity of 7.5 arcmin. The other had a disparity that varied according to a staircase procedure to obtain a PSE for depth. Observers were required to report which of the two images appeared to have greater depth. Burge et al. (2005) argued that:

> "…with a consistent standard display and an inconsistent comparison display subjects should require *more* disparity for the depth separation in the comparison display to appear identical to that in the standard (PSE > 7.5 arcmin). In contrast, with an inconsistent standard and a consistent comparison, *less* disparity should be required for depth separation in the comparison to appear identical to that in the standard (PSE < 7.5 arcmin). If this result is observed, we will have shown a quantitative effect of configural
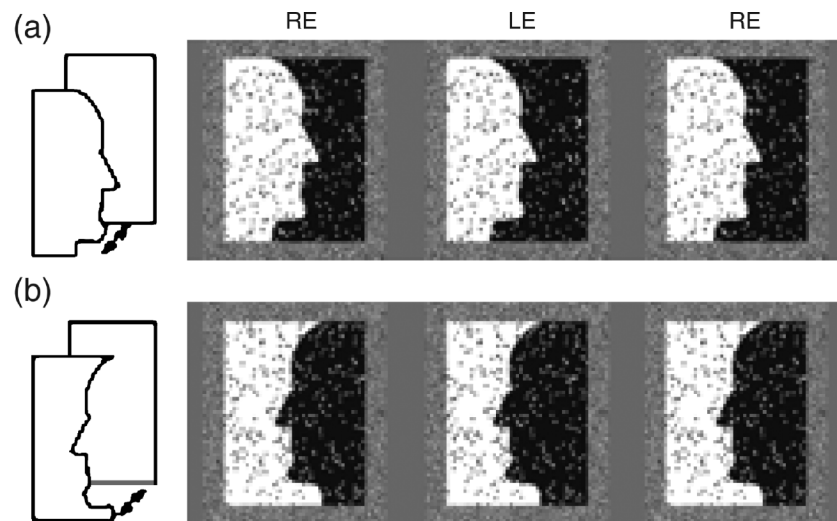
Figure 1. The figure used by Burge et al. (2005). We used the same figure but without the gray surround and with dot colors and viewing conditions altered (see text).

cues on metric depth perception, suggesting that the face side appears slightly closer due to its configural properties." p. 536

This was indeed the result obtained. In a second experiment, the consistent and inconsistent stimuli were each paired with stimuli that were neutral with respect to configural properties and PSEs were obtained by the same method. Burge et al. (2005) again reported that the face near stimulus was perceived as having greater depth than the face far stimulus using the criteria outlined above. This result was replicated by Bertamini, Martinovic, and Wuerger (2008) using the same method but with very different disparity differences (1.1, 1.6, and 2.6 arcmin) between figure and ground (face/non-face difference), and a significantly shorter viewing distance (2 meters). They also showed that the effect occurs when luminance contours are eliminated and the shape of figure and ground are given purely stereoscopically.

The intriguing result in the Burge et al. (2005) paper raises two issues. The first is that that the authors do not say why a face shape specified as stereoscopically closer than a background should appear to have a larger depth interval than a neutral shape with the same disparity relationship to the background. It would not seem ecologically advantageous for the face to elicit an exaggerated depth interval rather than that specified by disparity, which would normally be correct. On the other hand, it does seem plausible that a background with a face shape may provide some conflict with the stereoscopic cues to surface order and this may interfere with the perceived relative depth of the two surfaces. The second and related issue is that the forced choice method used by Burge et al. (2005) leaves open the possibility that the differences in depth reported by observers arose from response bias. In both experiments observers were asked

on each trial to say which of two displays has greater depth relative to its background: the stimulus with the face in front, or the stimulus with the non-face in front. It is possible that when observers are confronted with a forced choice of this kind, they are slightly biased to report faces as nearer when the decision is difficult, even if the metric depth experienced in these displays was unaltered by the configural face cues. The possibility that the effect is determined during the process of response selection rather than depth determination was raised by Bertamini et al. (2008) in discussing their similar results, but remained unresolved as they had employed the same methods as Burge et al. (2005). A bias to report face stimuli nearer, perhaps due to attention within the bipartite display, accounts for the similar and opposite effects found by Burge et al. (2005) in conditions of cue agreement and cue conflict. This symmetry in results between two very different stimulus situations is difficult to account for with a cue combination approach and Burge et al. (2005) do not attempt to account for it.

In order to determine whether these results were due to response bias, or due to a genuine distortion of perceived depth, a method is needed to determine how depth is perceived in these displays. To accomplish this goal, in our experiment depth of the foreground/background difference was measured for the two stimuli (face in front and non-face in front) using a depth matching method. Under these conditions we found no evidence for a distortion of perceived metric depth by ordinal depth cues.

Stimuli

We used the same basic stimulus used by Burge et al. (2005); see Figure 1, with some minor differences in

color, stimulus scale and viewing conditions. The outer dimensions of our stimuli (edges of the random dot surround) was 5.3×5.1 deg. centered in the plane of the monitor screen (which was 23 deg $\times$ 18 deg) and viewed from a distance of 85 cm. We substituted white dots on black and black dots on white for the black on red and red on black dots used by Burge et al. (2005), since their motivation for using red displays was to limit cross-talk in shutter glasses (which we did not use). Dot density was 18 per square deg. at the distance we used. The face in front and face behind conditions were each black on white on half the trials and white on black on the other half. Each of these conditions was presented equally often with the near face on the left and on the right. Our stimuli, with the left and the right eye's views presented on separate screens, were superimposed by means of an arrangement of mirrors forming a Wheatstone stereo-scope. The edges of the monitor screen were not visible.

The near surface (face or non-face) was more distant stereoscopically than the random dot surround by a disparity of 7 arcmin. We varied the disparity difference between the near and far surface by keeping the near surface constant and varying the disparity of the far surface. We used three disparities; 5.4, 8.7, and 12 arcmin (corresponding to simulated depths of 1.9, 3.0, and 4.1 cm, respectively). For each disparity we varied whether the nearer surface was face or non-face and was on the left or right side.

To avoid the complications of possible disparity match-ing, observers were not asked to match component depths directly, but rather, to match the depth difference between the foreground and background surface of the test stimulus by adjusting the depth difference of a comparison stimulus placed 1.6 cm (2.1 arcmin) below the test stimulus. The matching stimulus consisted of a white disc sprinkled with black random dots (1.8 cm/1.3 deg diameter) centered in a black square sprinkled with white random dots (side of 4.3 cm/2.9 deg). Observers adjusted the disparity of the white disc so that the depth difference between it and the black square appeared to match the depth difference between the face and non-face.

We added a pedestal disparity to the matching pattern so that it was stereoscopically at a different distance from the interval to be matched. Note that the perceived depth between two fixed disparities scales with distance: A disparity difference between two near surfaces will appear as a smaller depth than the same disparity difference seen at a greater distance. Thus, this method required observers to attend to the perceived depth difference between the two surfaces, and matches could not be achieved by simply matching disparities that were in the test images. The pedestal was achieved by placing the constant background of the matching stimulus (the black square) 3.5 arcmin disparity nearer than the nearer of the two surfaces to be matched. This meant that it was also 3.5 arcmin behind the random dot surround.

Each condition was replicated 6 times for each observer with all conditions presented in a randomized order.

Observers

There were 18 naive observers from the first year subject pool at UNSW and two further research assistants who were completely naive concerning the rationale of the experiment ($N = 20$). All were screened for stereopsis with the Titmus test as a criterion for participation.

Results

Two observers were removed from the analysis because of very large standard deviations in their settings (above 10 arcmin) in some conditions. A planned comparison ANOVA for repeated measures was carried out on the results of the remaining 18 observers. The means are shown in Figure 2. The main effect of disparity was highly significant ($F_{1,16} = 219.5$). The main effect of face/non-face was not significant ($F_{1,16} = .285$). The interaction of the linear and quadratic components of disparity with face/non-face was not significant ($F_{1,16} = .352$ and .947 respectively). Left/right position of face/non-face was not significant ($F_{1,16} = .285$) nor were any interactions involving this factor. The F critical ($p < 0.05$) for all these effects was 6.12. The mean difference between face-near/non-face far and non-face near/face far was .110 minutes of arcmin in favor of the former condition with 95% confidence intervals of $-.326$ and $+0.547$ arcmin (taken from an Individual T analysis based on
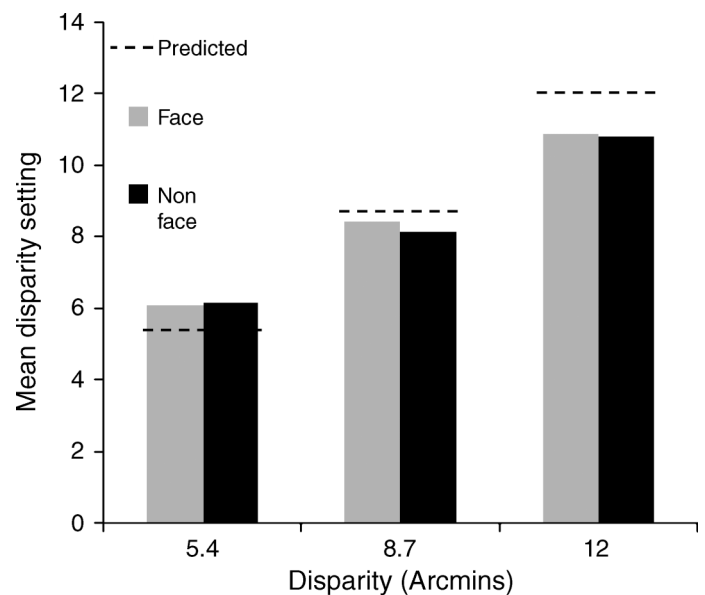

Figure 2. Bar graphs showing the disparity settings of the double probe for face near and non-face near conditions for three different disparities of face and non-face surfaces. Dotted lines show veridical matches. Error bars are not shown since the data are repeated measures. (See Cumming & Finch, 2005).

the ANOVA). These confidence intervals include zero and do not include the mean configural advantages in arcmin for face in front conditions found by Burge et al. (2005) for which the values ranged from .7 to 1.6 minutes arc.

We also carried out analyses of individual data. Ten observers had a higher mean setting for face in front and 8 for non-face in front. Since there were no significant differences between conditions with the face on the left and the face on the right these conditions were combined to make 12 replications per condition for each observer and an individual planned comparison ANOVA was carried out for each observer. Only two of the 18 observers had a significant main effect for face/non-face ($Fcrit_{1,11} = 6.72$), and these two had mean differences in opposite directions. Only one observer had a significant interaction between face/non-face and disparity level. The results of the individual observers therefore support the analysis based on their mean results. Thus, when observers are asked to explicitly judge depth intervals, we obtained no support for the hypothesis that there is greater depth between figure and ground when the face is placed stereoscopically closer and the non-face further, than for the opposite depth pattern. It should be noted that the precision of our results as measured by confidence intervals in minutes of arc was greater than in the Burge et al. (2005) study in which confidence intervals for the configural effect ranged from 1.48 and 2.08 minutes of arc in Experiment 1 with similar values (all above 1 minute arc) in Experiment 2. So although one can never prove a null result, it cannot be argued that our experiment lacked the precision or power to reveal an effect as large as Burge et al.'s (2005) if one were there.

Discussion

Bertamini et al. (2008) and Burge et al. (2005) reported that a configural cue can have an impact on the way observers respond to a forced choice paradigm where they are required to choose which of two patterns has the most depth. Such paradigms are incapable of discriminating between a difference that arises from a response bias and those that arise from a genuine distortion of perceived depth. In our experiment, observers were required to match the perceived depth interval of the displays used in Burge et al.'s (2005) studies. In this paradigm, we observed no difference in the settings of metric depth depending on which region appears as figure. The precision of our measurements was at least as high as those of Burge et al. (2005), so this failure cannot be attributed to a difference in statistical power in our studies. Our results indicate that the Burge et al. (2005) configural effect is likely to have arisen from response bias.

One difference between our experiment and Burge et al.'s (2005) was the viewing distance of the observers. They used a viewing distance of 3.25 meters. With a

disparity between the near and far surface of 7.5 arcmins, this is consistent with a depth of 40 cm between the two surfaces. We used a range of disparities between the two surfaces (5.4 to 12 arcmins) but our viewing distance was only 85 cm and the maximum predicted depth difference was thus only 4.2 cm. A reviewer raised the possibility that disparity may receive less weight at greater distances because disparity scales with distance (decreasing as the square of the distance for a given depth). However our disparities were not greater than those of Burge et al. (2005) so there is no reason to attribute our results to the nearer distance we used. Threshold depths expressed as disparities do not vary with distance within the relevant range (Bradshaw & Glennerster, 2006). Also given the very large perceived depth in the Burge et al. (2005) stimuli, we might expect that if a constant amount of extra depth were assumed to be provided by configural cues, it should be more difficult to detect (by Weber's Law) in their conditions than ours. The effects reported by Bertamini et al. (2008) at an intermediate distance with larger sizes and smaller disparities than Burge et al. (2005) indicate that the configural effect is not specific to a narrow set of parameters and Burge et al. (2005) do not state that it is. We think it likely that the Bertamini et al. results are also determined at the response selection level, which was also suggested by these authors.

Although our data do not support a view that configural face cues in a figure-ground arrangement influence metric depth, we do not claim that there are no contexts in which ordinal factors influence metric depth. On the contrary, we believe that there are cases where they do. These seem to fall into two categories. On the one hand, there is some evidence that interposition, an ordinal cue, can interfere with the perception of depth when in conflict with stereopsis (Schriever, 1925). Such conflicts are probably rare however when viewing natural scenes. In the other and more common category, ordinal cues serve essentially a veto function. They can provide information used by the visual system to indicate whether metric information is applicable. For example, Gillam and Cook (2001) showed that whether a cyclopean trapezoid emerging from a random dot stereogram was in front or behind the surround had a strong influence on whether its trapezoidal shape influenced the perceived metric stereoscopic slant to a disparity gradient across its surface. In the behind case, the trapezoid shape is attributed to the aperture through which it is viewed, rather than to the surface itself. In this arrangement, the perspective cue generated by the trapezoidal shape had little effect on the perceived slant of the stereoscopically defined surface within its boundaries, but it had a much more significant influence on perceived slant when the same surface was placed in front of the surround. In this latter configuration, the trapezoidal shape is intrinsic to the surface it bounds, and hence provides information relevant to the surface's slant, which is not true when the ordinal depth is reversed. Likewise, the introduction of a surface in the correct position for a

possible partial occlusion can change the stereo response to a disparate rectangle from strong slant to little slant (Häkkinen & Nyman, 1997). In these cases, an ordinal cue is informative about the contours to which it is appropriate to apply metric information provided by disparity.

There have also been a variety of other experiments that have shown that non-metric information can have a dramatic effect on the use of metric information. Meng and Sedgwick (2001) showed that the perception of the metric properties of surface layouts could be strongly affected by whether surfaces were perceived to be in contact. A similar effect of the role of contact relationships affecting perceived metric depth was reported by Kersten, Mamassian, and Knill (1997), who showed that the perceived motion trajectory and depth of a moving object could be dramatically altered by the presence of a shadow that either caused an object to appear to travel along a surface, or along a path off of the surface.

These associations between ordinal cues, or qualitative contact relationships, and their influence on metric depth are undoubtedly built in by evolution or learned. Such processes would presumably function to reinforce correlations between ecologically relevant cues to depth, and actual depth intervals. Burge et al. (2005) try and reconcile their finding of an ordinal effect on metric depth with Landy et al.'s (1995) requirement that depth cues be in the same units for combination. They speculate that an ordinal cue can acquire metric status by suggesting that an occluding surface could lead to an internalization of "the statistical likelihood of a metric depth value given that geometrically ordinal depth cue" (Burge et al., 2005, p. 541). In other words, they suggest that an ordinal cue could become associated with the most likely depth interval associated with that particular cue. Even if this were true, it seems unclear why faces should be assigned a different value than an arbitrary occluding surface, which would be needed to explain their results within the cue combination approach that they offer. We believe the more parsimonious account of their results, given the results reported herein, is that they reflect a form of response bias.

In conclusion, our results indicated that the effect of a figure-ground cue on perceived metric depth as demonstrated by Burge et al. (2005) did not occur when perceived depth in these displays was directly measured. The configural effect on metric depth appears to be a consequence of a forced-choice method that cannot distinguish between differences that arise from genuine transformations in perceived depth, and those that arise from response bias.

Corresponding author: Barbara Gillam.
Email: b.gillam@unsw.edu.au.
Address: School of Psychology, Mathews Building, University of New South Wales, Sydney, NSW 2052, Australia.

References

Bertamini, M., Martinovic, J., & Wuerger, S. M. (2008). Integration of ordinal and metric cues in depth processing. *Journal of Vision, 8*(2):10, 1–12, http://journalofvision.org/8/2/10/, doi:10.1167/8.2.10. [PubMed] [Article]

Bradshaw, M. F., & Glennerster, A. (2006). Stereoscopic acuity and observation distance. *Spatial Vision, 19,* 21–36. [PubMed]

Burge, J., Peterson, M. A., & Palmer, S. E. (2005). Ordinal configural cues combine with metric disparity in depth perception. *Journal of Vision, 5*(6):5, 534–542, http://journalofvision.org/5/6/5/, doi:10.1167/5.6.5. [PubMed] [Article]

Cumming, G., & Finch, S., (2005). Inference by eye: Confidence intervals and how to read pictures of data *American Psychologist, 60,* 170–180. [PubMed]

Gillam, B. J., & Cook, M. L. (2001). Perspective based on stereopsis and occlusion. *Psychological Science, 12,* 424–429. [PubMed]

Häkkinen, J., & Nyman, G. (1997). Occlusion constraints and stereoscopic slant. *Perception, 26,* 29–38. [PubMed]

Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception, 26,* 171–192. [PubMed]

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412. [PubMed]

Meng, J. C., & Sedgwick, H. A. (2001). Distance perception mediated through nested contact relations among surfaces. *Perception & Psychophysics, 63,* 1–15. [PubMed]

Peterson, M. A., & Gibson, B. S. (1994). Object recognition contributes to figure-ground organization: Operations on outlines and subjective contours. *Perception & Psychophysics, 56,* 551–564. [PubMed]

Schriever, W. (1925). Experimentelle studien über stereoskopisches sehen. *Zeitschrift für Psychology und Physiologie der Sinnesorgane, 96,* 113–170.