Multistability of overlapped face stimuli is dependent upon orientation

Isabelle Boutet, Avi Chaudhuri
Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montréal, Québec H3A 1B1, Canada; e-mail: isabou@ego.psych.mcgill.ca; http://www.psych.mcgill.ca/labs/cvl/
Received 15 July 2000, in revised form 6 November 2000

Abstract. Stimuli composed of two overlapped faces, one rotated 45° clockwise and the other 45° counterclockwise, produce perceptual rivalry whereby both faces cannot be simultaneously perceived. We obtained subjective and quantitative measures of this rivalry effect and examined if it persists with inverted stimuli. Our results show that upright stimuli are multistable, with alternations occurring from one face to the other within 2 s. Inverted stimuli were instead perceived as ambiguous in half of the trials, indicating weaker perceptual rivalry in that condition. We suggest that overlapped faces produce perceptual rivalry because each face is readily interpreted into a Gestalt, an effect that in turn is dependent upon orientation.

1 Introduction
We report a novel multistable phenomenon produced by stimuli composed of two overlapped faces, one rotated 45° clockwise and the other 45° counterclockwise (figure 1). The stimuli produced perceptual rivalry in that both faces could not be simultaneously perceived. Rather, only one of the faces remained perceptually dominant before alternating to the other face. This phenomenon displays similar characteristics as binocular rivalry effects (e.g. Engel 1956; Lumer et al. 1998; Yu and Blake 1992) and reversible figures (e.g. Necker cube, Rubin’s reversible goblet, young girl–old woman figure, etc; Attneave 1971) in that (i) at any one moment only one face can be clearly perceived, and (ii) the different percepts alternate periodically.

Figure 1. An example of the upright and inverted stimuli used in this experiment.

A possible determinant of this multistable effect is the propensity for faces to be perceived as holistic configurations. A number of studies have shown that configural properties are more important than featural properties during both the perceptual (Farah et al. 1995; Sergent 1984) and recognition (Tanaka and Sengco 1997) stages of face processing. There is convincing evidence that perception of a face as a holistic configuration is sensitive to alterations in orientation such that inverted faces may
instead be perceived as a collection of segregated facial features (Farah et al 1995; Kemp et al 1990; Rhodes et al 1993; Sergent 1984; Tanaka and Sengco 1997; Young et al 1987). Inverting the stimulus thus provides a test for the possibility that perception of a face as a holistic configuration contributes to the rivalry effect that we observed. (1)

The experiment reported in this paper served two purposes. First, to provide a measure of the rivalry effect, and, second, to examine if the effect persists when the orientation is changed. To lend experimental support to the observation that overlapped-face stimuli are multistable, we devised a subjective and quantitative measure of the effect using a sequential matching paradigm. After presentation of each overlapped stimulus, observers were asked to report their subjective experience of the stimulus in a forced-choice task and then to identify both faces among distractors in a recognition task. Overlapped stimuli were presented for 1, 2, or 3 s.

We reasoned that, if the two faces in a stimulus cannot be simultaneously perceived, then short presentation times should allow for the perception of only one of them, while longer presentation times should allow for alternations from one face to the next. Conversely, if there is no rivalry between the two faces, then both should be perceived at all presentation times, or the stimulus could be perceived as ambiguous with neither face being clearly perceived. It is assumed that participants made an attempt to perceive both faces within the presentation time allocated because they knew that recognition of both faces would be tested. As such, perception of only one of the two faces should reflect an inability to perceive both faces simultaneously rather than a tendency to attend to only one of the two faces.

Following this rationale, each trial was classified on the basis of whether one face (single), both faces (double), or neither face (no) had been clearly perceived and accurately recognised. Our prediction was that perceptual rivalry should produce a greater proportion of single than double trials for short presentation times, while the proportion of double trials would be greater than, or equivalent to, that of single trials for longer times because alternations from one face to the next might occur. An absence of rivalry should produce a greater proportion of double than single trials for all presentation times. Finally, it was predicted that ambiguous perception of the stimulus at any given presentation time would result in a greater proportion of no trials than single and double trials.

2 Methods
2.1 Participants
Twenty-three female and seven male students from McGill University participated in this study. Their ages ranged from 17 to 25 years. All participants had normal or corrected-to-normal vision.

2.2 Stimulus materials
120 digitised photographs of male faces were obtained from a face database at the University of Essex (http://hp1.essex.ac.uk/projects/vision/allfaces/). The original full-colour face images were converted to a 256 gray-level format. The face images were 200 × 200 pixels (subtending 7 deg × 7 deg at a viewing distance of 57 cm). All images were unknown to the subjects. They were in full-face view with a homogenous gray-level background. All images were equalised to an average luminance of 47.2 cd m⁻². Half of the faces were tilted clockwise (CW) and the other half counterclockwise (CCW) by 45°. Each CW face was randomly paired with a CCW face and overlapped with 50% transparency by means of the Adobe Photoshop 5.0 software. Half of the 60 stimuli created were inverted. Half of the observers were tested with these stimuli and the other half were

(1) The terms ‘upright’ and ‘inverted’ are used in this paper to refer to overlapped faces that are actually 45° off the cardinal axes.
tested with the same CW face–CCW face pairs but with the CW face rotated CCW and vice versa, and with the inverted stimuli shown upright and vice versa.

Participants were tested individually with the use of a Macintosh G3/266 computer. The stimuli were presented on a 21-inch Sony colour monitor. The screen was calibrated to linearised luminance values with an Optikon Universal Photometer. A neutral gray background of 18.6 cd m$^{-2}$ filled the screen.

2.3 Procedure
Participants were instructed that the goal of the experiment was to measure their subjective perception of the overlapped stimuli and to measure their ability to recognise the faces shown during encoding. No suggestion was given as to the possibility that they might experience a rivalry effect. Participants were tested on 30 trials, each consisting of the presentation of one overlapped stimulus, followed by a subjective measure, and then by a quantitative measure of the rivalry effect (see figure 2). The intertrial-time was 1 s.

2.3.1 Encoding. For each observer, 15 upright and 15 inverted stimuli were randomly chosen and presented during encoding. Each stimulus was presented for either 1, 2, or 3 s. The order of presentation of the stimuli and times were randomised.

2.3.2 Subjective measure—perceived saliency. The following text was written on the centre of the screen 1 s after the disappearance of the overlapped stimulus:

“Press ‘1’ if you clearly perceived both the CW and CCW images as whole, visible and independent entities. Press ‘2’ if you could only perceive the CW image as a whole, visible and independent entity. Press ‘3’ if you could only perceive the CCW image as a whole, visible and independent entity. Press ‘4’ if the stimulus looked like a scrambled mix of all the features that made up the two images, i.e. if neither image was perceived as a whole, visible and independent entity.”

Observers were instructed to press the key that best corresponded to their perceptual experience.

2.3.3 Quantitative measure—recognition accuracy. 1 s after a key press was recorded, two rows of four images were shown. One row contained CW images and the other CCW images (counterbalanced). The target faces were shown in the same size, orientation (ie, upright or inverted), and rotation (ie CW or CCW) as during encoding.
The six distractor faces were randomly chosen from those stimuli that were never shown during encoding. Participants had to press the key corresponding to the image that was shown during encoding for each row in the order of their choice.

3 Results

3.1 Subjective measure

A subjective measure was obtained by calculating the proportion of trials where both faces, a CW face, a CCW face, or neither faces were clearly perceived (figure 3). Data for two participants were discarded because they reported being able to clearly perceive both the CW and CCW images as whole, visible, and independent entities on all the trials for both the upright and inverted conditions. However, these participants showed poor recognition performance during quantitative assessment, suggesting that they had difficulties interpreting the descriptions provided.

Perceptual rivalry was evaluated by comparing the proportion of trials where either a single face (single perception—SP) or both faces (double perception—DP) were clearly perceived as indicated by the participant’s choice on the subjective measure. Our prediction was that rivalry would produce a greater proportion of SP than DP trials for short presentation times, while the proportion of DP trials would be greater than, or equivalent to, that of SP trials for longer times because an alternation from one face to the other is more likely. As figure 3 shows, a short presentation time of upright stimuli produced a greater proportion of SP than DP responses, whereas longer presentation times (ie 2 and 3 s) showed a marked decrease in SP responses at the advantage of DP responses. This was not the case for inverted stimuli, where the perception of neither face (no perception—NP) prevailed at all presentation times.

A $2 \times 3 \times 2$ repeated-measures ANOVA with stimulus orientation (upright and inverted), presentation time (1 s, 2 s, and 3 s), and trial type (DP and SP) as variables was performed on the average proportion of SP and DP trials obtained across participants. The main effects of orientation ($F_{1,27} = 59.14$, $p < 0.01$), presentation time ($F_{2,54} = 3.38$, $p = 0.04$), and trial type ($F_{1,27} = 9.26$, $p < 0.01$) were significant. The orientation $\times$ trial type ($F_{1,27} = 4.71$, $p = 0.04$), presentation time $\times$ trial type ($F_{2,54} = 6.09$, $p < 0.01$), and orientation $\times$ presentation time $\times$ trial type ($F_{2,54} = 5.09$, $p < 0.01$) interactions were also significant. The orientation $\times$ presentation time interaction was not significant ($F_{2,54} = 0.30$, $p = 0.75$).
Planned comparisons were used to evaluate perceptual rivalry for each presentation time and for upright and inverted stimuli separately. Upright stimuli presented for 1 s yielded more frequent SP than DP trials ($F_{1,54} = 25.89, p < 0.01$), but not stimuli presented for 2 s ($F_{1,54} = 0.04, p = 0.84$). Stimuli presented for 3 s yielded more frequent DP than SP trials ($F_{1,54} = 1.14, p = 0.05$). These results suggest that upright stimuli presented for 1 s induced rivalry in that only one of the two faces was clearly perceived. This finding provides experimental support for the rivalry effect we had casually observed. The difference between SP and DP trials disappeared after 2 s, suggesting that longer inspection times allowed alternations from one face to the next to occur. Inverted stimuli presented for 1 s ($F_{1,54} = 14.18, p < 0.01$), 2 s ($F_{1,54} = 10.60, p < 0.01$), and 3 s ($F_{1,54} = 8.12, p < 0.01$) yielded more frequent SP than DP trials. While these results indicate that the inverted stimuli produced perceptual rivalry, further inspection of the subjective reports points to a different conclusion.

We also examined the proportion of trials where participants reported that neither face was clearly perceived (no perception—NP). A $2 \times 3$ repeated-measures ANOVA with orientation (upright and inverted) and presentation time (1 s, 2 s, and 3 s) as variables was performed on the average proportion of NP trials. The main effects of orientation ($F_{1,27} = 65.23, p < 0.01$) and presentation time ($F_{2,54} = 7.55, p < 0.01$) were significant. The orientation $\times$ presentation time interaction was not significant ($F_{2,54} = 0.49, p = 0.61$). These results indicate that inverted stimuli yielded a greater proportion of NP trials (50%) than the upright stimuli (18%). This finding supports the notion that the inverted and upright stimuli induced a different perceptual experience. It can be appreciated from figure 1 that while one of the faces immediately pops out in the upright stimulus, the two faces appear to blend together in the inverted stimulus where some effort is required to perceive one of them as a whole configuration. Our interpretation is that segregation of one face from the other was much more difficult with the inverted stimuli than with the upright ones, explaining why neither face was clearly perceived in half the trials in the inverted case.

3.2 Quantitative measure

A quantitative measure was obtained by calculating the proportion of trials where both faces, a single face (CW or CCW), or neither face were accurately recognised. As can be seen in figure 4, single recognition (SR) for upright stimuli is greater than double recognition (DR) at 1 s duration, whereas at longer times the difference diminishes (2 s) and actually becomes negligible (3 s). With inverted stimuli, the proportion of SR trials always exceeded the DR trials for all presentation times. More importantly, the number of trials where neither face was accurately recognised (NR) was always greater than for the upright condition.

![Figure 4](image-url)  
**Figure 4.** Average percentage of trials obtained for the quantitative measure with the upright and inverted stimuli (30 subjects). Trials where both faces (DR), only one face (SR), or neither face (NR) were accurately recognised are illustrated. Error bars represent ±1 SE.
A $2 \times 3 \times 2$ repeated-measures ANOVA with stimulus orientation (upright and inverted), presentation time (1 s, 2 s, and 3 s), and trial type (DR and SR) as variables was performed on the average proportion of DR and SR trials obtained across observers. The main effects of orientation ($F_{1,29} = 43.26$, $p < 0.01$), presentation time ($F_{2,58} = 3.28$, $p = 0.05$), and trial type ($F_{1,29} = 41.51$, $p < 0.01$) were significant. The orientation $\times$ trial type interaction was almost significant ($F_{1,29} = 3.05$, $p = 0.09$). The presentation time $\times$ trial type interaction was significant ($F_{2,58} = 3.28$, $p = 0.05$). The orientation $\times$ presentation time ($F_{2,58} = 1.46$, $p = 0.24$) and orientation $\times$ presentation time $\times$ trial type ($F_{2,58} = 1.75$, $p = 0.18$) interactions were not significant. These results are not entirely consistent with the subjective measure since the orientation $\times$ presentation time $\times$ trial type interaction was not significant. We discuss this discrepancy later in section 3.3.

Planned comparisons were used to evaluate perceptual rivalry for each presentation time and for upright and inverted stimuli separately. Upright stimuli presented for 1 s yielded more frequent SR than DR responses ($F_{1,58} = 12.83$, $p < 0.01$), but not stimuli presented for 2 s ($F_{1,58} = 1.30$, $p = 0.26$) and 3 s ($F_{1,58} = 0.24$, $p = 0.63$). In agreement with the subjective data, this finding supports the notion that upright stimuli produced rivalry with alternations occurring within 2 s. Inverted stimuli presented for 1 s ($F_{1,58} = 8.59$, $p < 0.01$), 2 s ($F_{1,58} = 5.57$, $p = 0.02$), and 3 s ($F_{1,58} = 6.79$, $p = 0.01$) yielded a greater proportion of SR than DR trials. Since these comparisons support the existence of a perceptual rivalry effect with upright stimuli, the quantitative measure is essentially in agreement with the subjective one.

Because DR trials do not differentiate between an alternation and an absence of rivalry, we also analysed the proportion of trials where neither face was accurately recognised. A $2 \times 3$ repeated-measures ANOVA with orientation (upright and inverted) and presentation time (1 s, 2 s, and 3 s) as variables was performed on the average proportion of trials where neither face was accurately recognised (no recognition—NR). The main effects of orientation ($F_{1,29} = 43.26$, $p < 0.01$) and presentation time ($F_{2,58} = 3.28$, $p = 0.05$) were significant. The orientation $\times$ presentation time interaction was not significant ($F_{2,58} = 1.46$, $p = 0.24$). Consistent with the subjective data, these results indicate that inverted stimuli yielded a greater proportion of NR trials (25%) than did the upright stimuli (11%). This finding supports the notion that inverted and upright stimuli induced different perceptual experiences.

### 3.3 Comparison between subjective and quantitative measures

The analyses indicate that the orientation $\times$ presentation time $\times$ trial type interaction was significant for the subjective measure but not for the quantitative one. One possibility for this discrepancy is that a face that was not clearly perceived by the observer could nonetheless have been sufficiently encoded to allow its recognition. To examine this possibility, we compared the subjective and quantitative data relating to those trials where neither face was clearly perceived (NP) or accurately recognised (NR). If faces that are not clearly perceived can nonetheless be accurately recognised, then NP trials should be more frequent than NR trials. A $2 \times 2 \times 3$ repeated-measures ANOVA with measure type (NP and NR trials), orientation (upright and inverted), and presentation time (1 s, 2 s, and 3 s) as variables was performed on the average proportion of NP and NR trials. Quantitative data from the two participants whose data were discarded in the subjective measure were not considered for this analysis. The main effect of interest, measure type, was significant ($F_{1,27} = 18.02$, $p < 0.01$) with NP trials (34%) being more frequent than NR trials (19%). This difference might explain inconsistencies between the two measures whereby participants can accurately recognise a face in a forced-choice procedure, even if they report that the face was not clearly perceived.
3.4 Overall recognition accuracy

For the upright stimuli, means (with standard errors in parenthesis) for percentage correct recognition for both CW and CCW faces for the 1 s, 2 s, and 3 s presentation times were 57.33 (2.99), 65.00 (3.07), and 68.67 (2.75), respectively. For the inverted stimuli, means were 45.67 (3.56), 52.00 (3.20), and 55.33 (2.79), respectively. The overall effect of inversion on face recognition was evaluated by comparing face recognition for the upright stimuli with that for the inverted stimuli. A $2 \times 3 \times 2$ repeated-measures ANOVA with orientation (upright and inverted), presentation time (1, 2, and 3 s), and angle of rotation (CW and CCW) as variables was performed on the recognition accuracy data. The main effect of interest to us, orientation, was significant ($F_{1,29} = 31.98, p < 0.01$) with an 11% accuracy difference between upright and inverted faces. The main effect of presentation time was also significant ($F_{12,58} = 7.73, p < 0.01$). None of the interactions were significant.

4 Discussion

Casual observation of two upright overlapped faces oriented 90° to each other reveals a perceptual rivalry effect where both faces cannot be simultaneously perceived. We obtained subjective and quantitative measures of this phenomenon with both upright and inverted stimuli. We hypothesised that rivalry would lead to the perception and accurate recognition of only one of the two faces for short presentation times while longer times should allow for alternations from one face to the next, and thus in the perception and recognition of both faces. Because trials where both faces are clearly perceived or recognised can reflect either an absence of rivalry or the presence of alternations, we also examined the proportion of trials where neither face was clearly perceived or accurately recognised.

4.1 Upright stimuli

Upright stimuli presented for 1 s produced perceptual rivalry as evidenced by the greater proportion of single than double perception and recognition trials. The effect was not apparent at 2 s and 3 s presentation times, suggesting that alternation from one face to the next can occur within 2 s. The visual system is known to segregate one object from another by grouping visual inputs into separate tokens that correspond to known representations. Grouping visual elements into a Gestalt can, however, be ambiguous when the tokens that form a stimulus share homogeneous properties, or when alternative groupings are equally plausible. Reversible figures illustrate how the presence of two possible groupings at the same location can produce multistability in that only one grouping is consciously perceived at any given time. Alternation and competition between two plausible groupings has been attributed to mutually exclusive inhibitory mechanisms. According to this model, a particular organisation remains dominant until its neural substrate reaches a critical level of satiation or fatigue, at which point the rival organisation becomes dominant until it too becomes sufficiently fatigued (Long and Olszweski 1999; Long and Toppino 1981; Toppino and Long 1987).

We speculate that one possible explanation for the rivalry effect that we observed may rest on the notion that an upright face is readily represented as a Gestalt (Farah et al 1995; Sergent 1984; Tanaka and Sengco 1997; Young et al 1987). As such, faces in our overlapped stimuli may be particularly susceptible to the visual system’s propensity to group visual inputs into known representations. Parsing of the visual input into two separable entities at the same location may therefore engage inhibitory processes similar to those implicated in reversible figures. Because each face is readily represented as a holistic configuration, the presence of both plausible groupings at the same location is apparent. However, according to the fatigue/satiation model, passive bottom–up
processes would be recruited to suppress one of the two neuronal representations and therefore allow only one of the faces to dominate the perceptual experience.

4.2 Inverted stimuli

In our study, inverted stimuli yielded a greater proportion of single than double perception and recognition trials for all the presentation times. While these findings can be taken as evidence that inverted stimuli produced perceptual rivalry, analysis of those trials where neither face was clearly perceived (NP) or accurately recognised (NR) showed that this was not the case. Indeed, inverted stimuli produced a greater proportion of such trials than the upright ones, suggesting the operation of different processes in these two conditions.

One possibility is that faces in the inverted stimuli were not readily encoded as two segregated entities. As a result, the inverted stimuli may have been initially perceived as an ambiguous collection of facial features that did not cohere into a holistic representation. Because the participants were aware that the stimuli contained two faces and that they would be tested on their recognition, we speculate that cognitive top–down processes were applied to solve the perceptual puzzle, whereby observers actively constructed a holistic representation of one face after the other (Pelton et al 1969). This interpretation is consistent with the notion that learning/decisional mechanisms are involved during the perception of reversible figures (Gregory 1970; Rock 1975). According to this view, alternations may be due to a cyclical process of hypothesis testing whereby the perceptual system vacillates between two equally acceptable solutions. This process implies the operation of active mechanisms such as, for example, the allocation of attention to a specific region of a reversible figure during its interpretation (Tsai and Kolbet 1985). Similarly, prior knowledge of faces may have served to segregate those features that pertained to a given face during perception of inverted stimuli to produce a coherent Gestalt.

4.3 Comparison to the classical inversion effect

We have reported a difference of 11% in recognition accuracy between the faces presented in the inverted and upright conditions. This effect is weaker than previously observed in studies where faces were presented in isolation with a 180° rotation from true upright to inverted (eg Diamond and Carey 1986, ~ 20%; Yin 1969, ~ 25%). The fact that the faces in our stimuli were rotated 45° from the cardinal axes, and as such were not truly upright or inverted, may be responsible for the weaker effect. However, the fact that the accuracy and subjective measures yielded statistically significant main effects of orientation suggests that inversion of our stimuli was sufficient to induce changes in configural information processing similar to those reported in previous studies (Farah et al 1995; Kemp et al 1990; Rhodes et al 1993; Sergent 1984; Tanaka and Sengco 1997; Young et al 1987).

4.4 Possible determinants of the multistable effect

Studies on binocular rivalry indicate that figures whose features can be organised into a coherent figure tend to predominate, and that sensory organisation of these figures is often dependent upon orientation. For example, it has been shown that an upright face or camouflaged Dalmatian figure dominates over their inverted versions (Engel 1956; Yu and Blake 1992). We have shown here that a similar perceptual rivalry effect occurs with facial stimuli whose organisation into a Gestalt is particularly sensitive to orientation.

If mechanisms similar to those involved in binocular rivalry are involved in the effect observed here, then our interpretation implies that faces in the upright condition should be organised into a holistic shape before the rivalry can be resolved. It has been shown that competition between binocularly rival complex figures is resolved...
before information reaches higher visual areas such as the inferior temporal area IT (Sheinberg and Logothetis 1997; Tong et al 1998). Our interpretation of the multistable face effect suggests that key aspects of configural face information are established in neural loci prior to those where resolution of the rivalry effect takes place.

The tendency for upright faces to be perceived as holistic configurations may be only partially responsible for the multistable effect. Other factors such as familiarity may have also contributed to the effect. Dominance of an upright face or Dalmatian figure over its inverted versions, for example, can be attributed to familiarity (Engel 1956; Yu and Blake 1992). In our stimuli, familiarity with upright faces may have facilitated figure-from-ground segregation. However, it is difficult to dissociate familiarity from holistic face encoding because sensory organisation of a face as a whole configuration has been attributed to expertise (Carey and Diamond 1994; Gauthier and Tarr 1997; Gauthier et al 1998; Tanaka and Gauthier 1997).

Interpretations based on attentional mechanisms have been proposed for other multistable effects (Attneave 1971; Lumer et al 1998). For instance, dominance of a given aspect of a reversible figure may be the result of competition between two possible interpretations (Attneave 1971). We initially speculated that upright stimuli produce perceptual rivalry because encoding of one face as a whole exhausts limited attentional resources. While a study by Reinitz et al (1994) indicates that holistic face encoding is attentionally demanding, we have recently found that perception of faces as holistic configurations requires little attention (Boutet et al 2000). The possible role of attention in the multistable phenomenon observed here is thus undetermined.

Attention may contribute to the perception of reversible figures by mediating ocular displacements towards a focal area that contains information important for the perception of the predominant interpretation (Garcia-Perez 1992; Ruggieri and Fernandez 1994; Scotto et al 1990; Tsal and Kolbet 1985). As such, perceptual rivalry could be the result of eye movements towards two different regions of the overlapped stimuli, one region being important for segregation of the CW face, and the other for segregation of the CCW face. While eye movements are likely to be involved in the perception of any reversible figure, they are not the sole determinant of perceptual rivalry because alternations during the perception of reversible figures can occur in the absence of eye movements (Attneave 1971).

Composite portraiture (Galton 1883) illustrates one example where overlapped faces do not produce multistability. Composite portraits are created by overlapping upright faces on a photographic film, the result of which is similar to what is obtained by overlapping transparencies in Photoshop software. The fact that non-slanted faces are not multistable and appear to be fused together suggests that rotating faces was a determinant of the rivalry effect that we have reported in this study. Overlapped slanted faces produce a striking difference from the non-slanted condition. With slanted faces, grouping visual features that pertain to each face produces two equally plausible solutions at the same location. This difference may explain why overlapped slanted faces produce rivalry while composite portraits do not.

4.5 Concluding remarks

We have suggested that face perception in our upright stimuli is determined by the tendency for faces to be represented as holistic configurations, inhibitory mechanisms are in turn recruited to suppress one representation to allow the alternative to dominate the perceptual experience. Because configural information is not readily constructed from inverted faces, face perception in our inverted stimuli was likely mediated by cognitive processes. Whether or not stimuli created by overlapping other complex objects would also produce rivalry remains to be determined. Given that configural properties are less important during the processing of objects for which expertise has
not been developed (Tanaka and Gauthier 1997; Tanaka and Sengco 1997), stimuli created by overlapping such objects may be perceived as an ambiguous collection of features and produce a similar effect as that observed with our inverted stimuli. Conversely, previous knowledge of the world could facilitate segregation of two overlapped objects and give rise to a multistable percept. Although we have discussed several possible determinants of the multistable face effect, it is evident that these remain at best speculative and that further study is required to identify the causes of this interesting phenomenon.

Acknowledgements. The study was supported by research grants to AC from the Medical Research Council (MRC) of Canada and the Natural Sciences and Engineering Research Council (NSERC) of Canada, and by a graduate scholarship to IB from NSERC. We thank Drs Fred Kingdom and Jim Ramsay, Cary Kogan, and two anonymous referees for their helpful comments.

References

Attneave F, 1971 “Multistability in perception” Scientific American 225(6) 63 – 71
Engel E, 1956 “The rôle of content in binocular resolution” American Journal of Psychology 69 87 – 91
Galton F, 1883 Inquiries into Human Faculty and Its Development (London: MacMillan)
Kemp R, McManus C, Pigott T, 1990 “Sensitivity to the displacement of facial features in negative and inverted images” Perception 19 531 – 543
Long G M, Olzwseski A D, 1999 “To reverse or not to reverse: When is an ambiguous figure not ambiguous?” American Journal of Psychology 112 41 – 71
Long G M, Toppino T C, 1981 “Multiple representations of the same reversible figure: implications for cognitive decisional interpretations” Perception & Psychophysics 10 231 – 234
Rock I, 1975 An Introduction to Perception (New York: MacMillan)
Sergent J, 1984 “An investigation into component and configural processes underlying face perception” British Journal of Psychology 75 221 – 242
Sheinberg D L, Logothetis N K, 1997 “The role of temporal cortical areas in perceptual organization” Proceedings of the National Academy of Sciences of the USA 94 3408 – 3413
Tanaka J, Gauthier I, 1997 “Expertise in object and face recognition” The Psychology of Learning and Motivation 36 83–125
Young A W, Hellawell D, Hay D C, 1987 “Configural information in face perception” Perception 16 747–759