Japanese Psychological Research
2005, Volume 47, No. 3, 216–222
Special Issue: Visual perception and perceptual processing in real and virtual environments

The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism[1]

ADRIAN SCHWANINGER[2]
Max Planck Institute for Biological Cybernetics, Tübingen 72076, Germany; and University of Zurich, Switzerland

FRED W. MAST
University of Zurich, Switzerland

Abstract: The effect of orientation on face recognition was explored by selectively altering facial components (eyes and mouth) or by changing configural information (distances between components). Regardless of the type of change, a linear increase in reaction time for same-different judgments was revealed when the faces were rotated away from upright. The analyses of error scores indicated that the detection of altered components was only slightly affected by orientation, while orientation had a detrimental effect on the detection of configural changes. These results are consistent with the assumption that rotated faces overtax an orientation normalization mechanism so that they have to be processed by mentally rotating parts, which makes it difficult to recover configural information.

Key words: face-inversion effect, mental rotation, component and configural processing, featural processing, face recognition.

It is already well known by painters and Gestalt psychologists that face processing is highly dependent on orientation (e.g., Köhler, 1940). Yin (1969) revealed that face recognition is disproportionately affected by inversion when compared with the recognition of other mono-oriented objects, such as airplanes, houses, and stick figures of men in motion. This finding has been referred to as the face-inversion effect. Subsequently, several studies have provided further evidence for the existence and robustness of this phenomenon (for reviews see Schwaninger, Carbon, & Leder, 2003; Valentine, 1988).

According to Farah, Drain, and Tanaka (1995) "face perception is holistic and the perception of holistically represented complex patterns is orientation sensitive" (p. 633). In this case holistic means that faces are processed and stored in memory as unparsed perceptual wholes, in which individual parts are not explicitly represented. According to the authors, such holistic processing is impaired when faces are substantially rotated away from their upright orientation, which results in the face-inversion effect (Farah et al., 1995; Tanaka & Farah, 1991; Tanaka & Farah, 1993).

An alternative explanation for the face-inversion effect is based on a qualitative distinction between component and configural information. The term component (or componential, piecemeal, featural) information has been used for facial elements that are perceived

---

[1] Adrian Schwaninger was supported by a grant from the European Commission (CogVis, IST-2000-29375). Fred Mast was supported by SNF grant no. 611-066052.

[2] Correspondence concerning this article should be sent to: Adrian Schwaninger, Max Planck Institute for Biological Cybernetics, Department Bülthoff, Spemannstr. 38, 72076 Tübingen, Germany. (Email: adrian.schwaninger@tuebingen.mpg.de)

as distinct parts of the whole, such as the eyes, mouth, nose or ears. The term configural information has been referred to as the "spatial interrelationship of facial features" (Bruce, 1988, p. 38). Similar meaning is conveyed by the terms configurational, spatial-relational, and second-order relational information. In practice, different manipulations have been used to change configural information, but one widely used method consists of altering the distance between components (Leder & Bruce, 1998; Murray, Yong, & Rhodes, 2000; Searcy & Bartlett, 1996; Sergent, 1984; Tanaka & Sengco, 1997).

According to the component configural hypothesis, processing configural information is strongly impaired when faces are turned upside-down. In contrast, processing component information should be relatively orientation invariant (e.g., Bartlett & Searcy, 1993; Carey & Diamond, 1977; Diamond & Carey, 1986; Searcy & Bartlett, 1996; Sergent, 1984; for a recent review see Schwaninger et al., 2003).

A third explanation for the face-inversion effect has been provided by Rock (1973, 1974, 1988). According to his view (mental-rotation hypothesis), complex stimuli such as faces overtax a mental-rotation mechanism when they are substantially rotated away from the upright. Rotated faces have to be processed by mentally rotating parts (or components) one after the other and this makes it difficult to recover holistic or configural information. Proponents of both the holistic and the component configural hypothesis have noted the explanatory power and have cited Rock's mental-rotation hypothesis. For example, Farah et al. (1995) have pointed out that the deeper answer to the question "Why is face recognition so orientation sensitive? … will concern capacity limitations of the orientation normalization process" (p. 633). Similarly, Searcy and Bartlett (1996) mentioned that the difficulty of processing configural information in disoriented faces could be due to the capacity limitations of a mental-rotation mechanism.

The main aim of this study was to test Rock's hypothesis directly. To this end, a sequential same-different matching task was used, in which selective changes of component or configural information had to be detected. If Rock was right, rotated faces can only be processed by mentally rotating parts (component information). As a consequence, detecting component changes, such as replaced eyes and mouth, would remain unaffected by rotation, whereas error scores would increase substantially when configural changes have to be detected in rotated faces. Moreover, because mentally rotating facial features takes time, it would be expected that reaction time (RT) increases with increasing rotation from the upright. This effect should be found in both tasks, that is, for the detection of component changes as well as for detecting configural changes.

Method

Participants

Sixty-four students of the University of Zurich volunteered as participants in this study. They were randomly assigned to one of two groups. In the first group, 16 men and 16 women had to detect component changes. The second group (16 male and 16 female participants) was tested in a condition in which configural changes had to be detected. All participants had normal or corrected-to-normal vision and were naïve as to the purpose of this study.

Materials

Stimuli were created from grayscale photographs of six people (three men and three women) who had agreed to be photographed and to have their pictures used in psychology experiments. The original grayscale pictures were front-facing and with a neutral expression. Digital images were obtained by developing the photographs on Kodak Photo CD™. These images were altered using image-processing software (Adobe Photoshop and Canvas). First, all images were scaled proportionally to have the same interpupillary distance. Then the hair was removed and the pictures were placed on a black background. These images constituted the set of six original images. Three anchor points for components were determined: the center of each pupil and the middle of the upper lip contour. The set of six faces with altered component

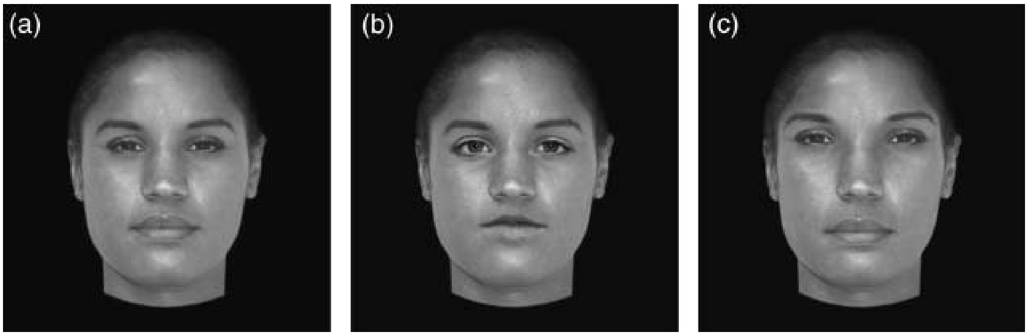

Figure 1. Example stimuli. (a) Original face, (b) component change, (c) configural change.

information was created by replacing the eyes and the mouth with components from another face of the same size. The location of new components was the same as in the original images (with an accuracy of 1 pixel concerning the anchor points defined above). New anchor points were determined in order to produce configural changes. The interpupillary distance, the distance between the pupils and the lower contour of the nose, and the distance between the nose and the mouth were scaled by constant factors (1.16, 1.14, and 1.23, respectively). The eyes and the mouth of the original images were then moved to the new anchor points. This resulted in empty skin areas that were filled with skin patches of the original images in order to ensure a selective change of configural information. All items were copied at seven different orientations (0°, 30°, 60°, 90°, 120°, 150°, 180°). Figure 1 contains examples of the stimuli.

Procedure

The experiments were conducted in a dimly lit room. Participants were seated in front of a computer monitor (17-in screen) at a distance of 0.48 m (1.6 feet). The stimuli covered 10° of visual angle and the viewing distance was maintained using a head rest. A sequential same-different matching task was used. A warning tone (one beep) started each trial. After 300 ms, an upright face was presented for 3000 ms followed by a 1000-ms blank. A warning tone (two beeps) announced the second face, which appeared after 300 ms in any one of seven clockwise rotated orientations 0° (upright), 30°, 60°, 90° (horizontal), 120°, 150°, 180° (upside-down). Whether the two faces were same or different had to be indicated by pressing a key (labeled “same” and “different”). Participants were instructed to respond as quickly and accurately as possible. Half the participants pressed the “same” key with their preferred hand and the others used the non-preferred hand. In the component condition, “different trials” consisted of faces with altered components (eyes and mouth). In the configural condition, “different trials” involved faces in which the configural information had been altered as explained above. Following the participant’s response, a 1000-ms blank field was displayed and the next trial started. Eight random orders were generated using the following constraints: (1) the same orientation was not repeated on consecutive trials; (2) the same face stimulus was not repeated on consecutive trials; and (3) there were no more than four consecutive “same trials” or “different trials.” The eight random orders were counterbalanced across the two conditions (component changes vs. configural changes), the sex of the participants and the assignment of the response buttons. There were 84 trials per experiment: 2 (same/different) × 6 (items) × 7 (orientations).

Prior to the experiment, a learning session was conducted. First, eight practice trials were carried out in order to familiarize the participants with the task. These stimuli were used in the practice trials only. Second, the six experimental pairs consisting of the original and the altered version were shown for 5 s each and the participant was instructed to memorize these pairs. The participants were not informed whether these pairs depicted faces of two different individuals or whether faces of the same individual had been manipulated. The purpose of this learning phase was to allow participants to form upright memory representations of the faces used in the experiment, thereby making the encoding conditions more similar to real-life situations. Third, 12 practice trials were carried out (six "same trials" and six "different trials") that contained the experimental face pairs presented sequentially in the upright orientation only. If the participant produced more than one error, these practice trials were repeated once (this occurred for only five of the 64 participants).

Results

Individual data were averaged across different faces in order to eliminate an item-specific factor. Separate and combined analyses were carried out on error scores of "different trials" and "same trials." Data were discarded if participants did not respond within 5 s. This occurred in only 0.13% of the trials (seven of the 5376 cases).

Analysis of error scores

Error scores of "different trials." A two-factor analysis of variance (ANOVA) with condition (detection of configural vs. component changes) as between-subjects factor and orientation as within-subjects factor was carried out[3] on error scores of "different trials." There were main effects of condition, $F(1,62) = 30.53$, $p < 0.001$, and orientation, $F(5,307) = 15.60$, $p < 0.001$, and there was an interaction between condition and orientation $F(5,307) = 11.03$, $p < 0.001$. As depicted in Figure 2, changes of orientation had a detrimental effect on the detection of configural manipulations, whereas the detection of component alterations was not affected by orientation. Separate one-factor within-subjects ANOVAs revealed that the effect of orientation on the detection of component changes did not reach statistical significance, $F(4,134) = 1.32$, while there was a strong main effect of orientation on the detection of configural alterations $F(4,137) = 17.01$, $p < 0.001$.

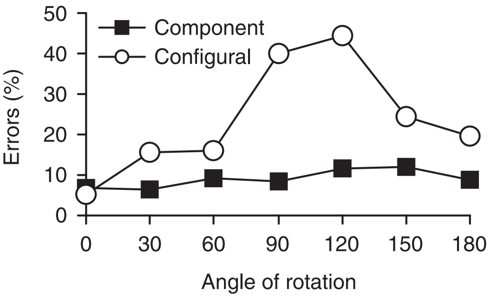

Figure 2. Mean error scores for "different trials" in the component condition (detection of component changes) and the configural condition (detection of configural changes).

Error scores of "same trials." A two-factor ANOVA with condition (component vs. configural changes) as between-subjects factor and orientation as within-subjects factor revealed a main effect of orientation, $F(4,261) = 24.78$, $p < 0.001$. There was no effect of condition, $F(1,62) = 1.52$, but there was an interaction between condition and orientation $F(4,261) = 2.46$, $p < 0.05$. Separate one-factor within-subjects ANOVAs showed a main effect of orientation for the component condition $F(4,115) = 18.59$, $p < 0.001$, as well as for the configural condition, $F(4,136) = 7.00$, $p < 0.001$. As depicted in Figure 3, the error scores of "same trials" increased with increasing rotation from the upright. This increase was even more pronounced for "same trials" in the component condition, thus yielding the significant interaction between condition and orientation.

Analysis of reaction times

Reaction times for "different trials." A two-factor ANOVA with condition (configural vs. component changes) as between-subjects factor

---

[3] In all analyses in this study, if Mauchly's (1940) test of sphericity showed a significant deviance ($\alpha = 0.05$) from equicorrelation for a repeated factor or for a combination of factors including at least one repeated factor, Greenhouse and Geisser's (1959) Epsilon was used to adjust the degrees of freedom for the averaged tests of significance.

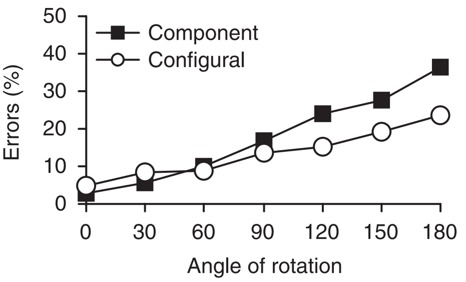

Figure 3. Mean error scores for "same trials" in the component condition (detection of component changes) and the configural condition (detection of configural changes).

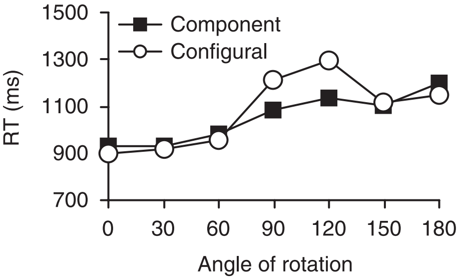

Figure 4. Mean correct reaction times (RT) for "different trials" in the component condition (detection of component changes) and the configural condition (detection of configural changes).

and orientation as within-subjects factor on correct RT of "different trials" revealed a main effect of orientation, $F(4,224) = 18.64$, $p < 0.001$. In contrast to the analysis of error scores, the analysis of RT gave no main effect of condition, $F(1,59) = 0.19$, and the interaction between condition and orientation was not significant, $F(4,224) = 1.61$. Separate one-factor ANOVAs on correct RT revealed a main effect of orientation for the detection of component changes $F(4,138) = 12.87$, $p < 0.001$, as well as for the detection of configural alterations $F(3,89) = 8.61$, $p < 0.001$ (Figure 4).

Reaction times for "same trials." A two-factor ANOVA on correct RT of "same trials" revealed a main effect of orientation, $F(4,248) = 39.40$, $p < 0.001$. As for the error scores, there was no main effect of condition for RT, $F(1,60) = 2.33$. The interaction between condition and orientation was significant, $F(4,248) = 3.10$, $p < 0.05$. Separate one-factor within-subjects ANOVAs revealed a main effect of orientation for the component and the configural condition $F(4,117) = 26.16$, $p < 0.001$, and $F(4,108) = 15.27$, $p < 0.001$, respectively (Figure 5).

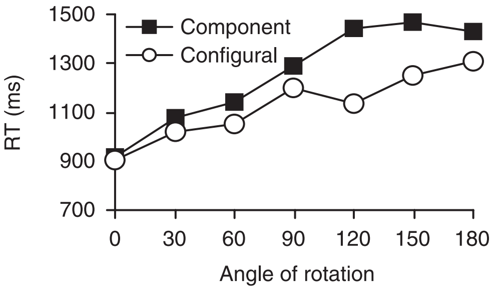

Figure 5. Mean correct reaction times (RT) for "same trials" in the component condition (detection of component changes) and the configural condition (detection of configural changes).

Discussion

The analyses of error scores of "different trials" revealed that orientation had no effect on error scores for detecting component changes, while the detection of configural alterations was strongly impaired when faces were substantially rotated away from the upright position. This result poses problems for a purely holistic view of face processing, which implies that rotating a face disrupts the processing of what is nominally component and configural information. A purely holistic view of face processing therefore fails to explain why error scores were highly affected by orientation when configural changes had to be detected, whereas detecting component changes remained orientation invariant. At the same time, the results supported the component-configural hypothesis as well as the mental-rotation hypothesis. They both predict strong impairment by rotation for the detection of configural alterations, while the detection of component changes should remain relatively unaffected. Note, however, that only the mental-rotation hypothesis explicitly predicts an increase in response time with increasing angle of rotation in both conditions (detection of component and configural changes). Because faces are so complex, they overtax an orientation normalization mechanism and rotated faces can only be processed by mentally rotating parts (component information). This takes more time the more a face is rotated from the upright and applies to both the component and configural condition. Indeed, in both conditions RT increased with increasing angular disparity following a similar linear trend.

However, there was a somewhat unexpected finding for the error scores of detecting configural changes. Instead of a monotonic increase, "different trials" in the configural condition showed that participants made the most errors at intermediate orientations of 90° and 120°, and not when the faces were presented upside-down. Interestingly, a similar effect has been found in object-naming studies. The time to name line drawings of natural objects has been found to increase linearly from upright to 120° of planar rotation, while naming times for 180° are often faster than those for 120° (e.g., Jolicoeur, 1985; Murray, 1995a, 1995b, 1997). However, such nonlinear effects are present primarily on the initial trials; after practice, they are usually diminished or even disappear. In fact, some studies suggest that when the stimulus set contains orientation-invariant information, the effects of orientation disappear following experience (Murray, 1999), which can occur even after a single presentation of objects in a block of trials (Murray, Jolicoeur, McMullen, & Ingleton, 1993). Interestingly, in our study, strong effects of orientation remained stable even after a remarkable amount of practice. This is consistent with the view that a transition to orientation-invariant processing could not take place and the subjects had to rely on normalization mechanisms for detecting facial alterations. An explanation for nonlinear effects of orientation has been provided by Corballis, Zbrodoff, Shetzer, and Butler (1978). They suggested that it might be possible to "mentally flip" an inverted picture out of the plane to match it to a memory representation (see also Koriat, Norman, & Kimchi, 1991). If it is assumed that mental flipping is possible when faces are inverted, one would expect that configural changes can be detected better at upside-down presentations than when faces are presented in intermediate orientations (see Figure 2).

The results obtained in "same trials" are also consistent with the mental-rotation hypothesis. In these trials no difference between the component and configural condition is expected because "same trials" always contain the same stimuli. According to the mental-rotation hypothesis, participants would mentally rotate parts in order to verify that the sequentially presented stimuli are indeed the same. This is true for the condition in which component changes had to be detected as well as for detecting configural changes. Because in both conditions "same trials" contained identical faces, no differences between conditions are expected. Indeed, there were no main effects of condition (detection of component vs. configural changes) for "same trials," neither in error scores nor in response times.

While the above-mentioned finding of nonlinear effects for processing configural information certainly requires additional investigation, several important theoretical contributions result from this study. First, the finding that component changes could be detected independent of orientation clearly indicates the existence of explicit part-based or component representations, whether they bear a hierarchical relation to whole face representations, or whether they constitute an independent population of representations. Moreover, our results suggest that when faces are rotated it is possible to process component information and mentally rotate facial features in order to match them to upright memory representations. Because mentally rotating a face as a whole overtaxes the orientation normalization mechanism, configural information is hard to recover and detecting configural changes becomes a very difficult task. Because face recognition relies strongly on detecting subtle configural differences between faces, a strong effect of inversion is observed. This might be the deeper answer to the question "Why is face recognition so orientation-sensitive?"

References

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, **25**, 281–316.

Bruce, V. (1988). *Recognizing faces*. Hillsdale, NJ: Erlbaum.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, **195**, 312–314.

Corballis, M. C., Zbrodoff, N. J., Shetzer, L. I., & Butler, P. B. (1978). Decisions about identity and orientation of rotated letters and digits. *Memory and Cognition*, **6**, 98–107.

Diamond, R., & Carey, S. (1986). Why faces are not special: an effect of expertise. *Journal of Experimental Psychology: General*, **115**, 107–117.

Farah, M. J., Drain, H. M., & Tanaka, J. W. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance* **21**, 628–634.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, **32**, 95–112.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, **13**, 289–303.

Köhler, W. (1940). *Dynamics in psychology*. New York: Liveright.

Koriat, A., Norman, J., & Kimchi, R. (1991). Recognition of rotated letters: Extracting invariance across successive and simultaneous stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 444–457.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology, A*, **5**, 449–473.

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Annals of Mathematical Statistics*, **11**, 204–210.

Murray, J. E. (1995a). Imagining and naming rotated natural objects. *Psychonomic Bulletin and Review*, **2**, 239–243.

Murray, J. E. (1995b). Negative priming by rotated objects. *Psychonomic Bulleting and Review*, **2**, 534–537.

Murray, J. E. (1997). Flipping and spinning: Spatial transformation procedures in the identification of rotated natural objects. *Memory and Cognition*, **25**, 96–105.

Murray, J. E. (1999). Orientation-specific effects in picture matching and naming. *Memory and Cognition*, **27**, 878–889.

Murray, J. E., Jolicoeur, P., McMullen, P. A., & Ingleton, M. (1993). Orientation-invariant transfer of training in the identification of rotated natural objects. *Memory and Cognition* **21**, 604–610.

Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science*, **11**, 492–496.

Rock, I. (1973). *Orientation and form*. San Diego, CA: Academic Press.

Rock, I. (1974). The perception of disoriented figures. *Scientific American*, **230**, 78–85.

Rock, I. (1988). On Thompson's inverted-face phenomenon [Research note]. *Perception*, **17**, 815–817.

Schwaninger, A., Carbon, C. C., & Leder, H. (2003). Expert face processing: Specialization and constraints. In G. Schwarzer & H. Leder (Eds.), *Development of face processing* (pp. 81–97). Göttingen: Hogrefe.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 904–915.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, **75**, 221–242.

Tanaka, J. W., & Farah, M. J. (1991). Second-order relational properties and the inversion effect: Testing a theory of face perception. *Perception and Psychophysics* **50**, 367–372.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, **46A**, 225–245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, **25**, 583–592.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**, 471–491.

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory and Cognition*, **16**, 556–566.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, **81**, 141–145.

(Received March 16, 2004; accepted March 26, 2005)