

## Holistic processing underlies gender judgments of faces

MINTAO ZHAO AND WILLIAM G. HAYWARD  
University of Hong Kong, Hong Kong, China

In three experiments, we investigated whether holistic processing underlies gender judgments about faces. Chinese participants were asked to make gender judgments for inverted, scrambled, or composite faces. Results showed that judgments were dramatically impaired by these manipulations (as compared with performance for normal upright faces), demonstrating three hallmark effects of holistic face processing that have been observed in perception of face identity. Whether the test faces were Chinese or Caucasian showed no effect on holistic processing of gender perception, in contrast to studies of identity analysis. These results suggest that holistic processing is a general mechanism for different aspects of face perception and are consistent with the idea that physiognomic properties that determine the gender of a face are universal, rather than race specific.

Much research has shown that people tend to perceive and recognize a face as a perceptual whole, rather than as a mere combination of different face components; this style of processing is referred to as *holistic face processing* (e.g., Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987; see Maurer, Le Grand, & Mondloch, 2002, for a review).<sup>1</sup> For example, recognizing familiar or newly learned faces is disproportionately impaired when holistic face processing is disrupted by inversion of a face (the *face inversion effect*; Yin, 1969), by presenting scrambled faces (Schwaninger, Lobmaier, & Collishaw, 2002) or isolated face parts (the *part-whole effect*; Tanaka & Farah, 1993), or by fusing the face parts of one person together with those of another (the *composite face effect*; Young et al., 1987).

Although these findings have frequently been cited as evidence supporting the idea that faces are processed holistically (e.g., Farah, Wilson, Drain, & Tanaka, 1998), they have been observed primarily in the processing of face identity. In contrast to the many studies demonstrating holistic processing of face identity, little has been done to investigate whether these holistic-processing effects can be generalized to other visual analysis routes for a face. The participants in the studies noted above were asked to either name the top or bottom half of a face (e.g., Young et al., 1987) or judge whether a face or an isolated face part belonged to a learned face identity (e.g., Tanaka & Farah, 1993; Yin, 1969). Processing of face identity, however, according to Bruce and Young's (1986) influential face recognition model, is only one visual analysis route for faces and runs parallel to others extracting facial expression, gender, and so forth. Therefore, the claim that faces are processed holistically is problematic without evidence from visual face-processing tasks other than identification.

In the present study, our aim was to test whether or not judgments of face gender would also be supported by the holistic-processing mechanism—specifically, whether or not those hallmark effects of holistic processing could be observed in a gender discrimination task. Determining whether holistic processing underlies face gender perception is crucially needed for an understanding of both its generality across different aspects of face perception and how the gender of a face might be typically coded. First, the holistic nature of face gender perception, which was assumed to proceed along a different route than identity and expression analysis (Bruce & Young, 1986), has not yet been clearly established. Therefore, we cannot rule out the possibility that holistic processing, which has been assumed to be a general mechanism underlying face perception, may be specific only to certain visual analysis routes (e.g., identification). Second, the role of holistic processing in making face gender decisions has been largely overlooked in the literature. Instead, most studies have emphasized the role of face features in the extracting of the gender of a face (e.g., Brown & Perrett, 1993; Dupuis-Roy, Fortin, Fiset, & Gosselin, 2009; Schyns, Bonnar, & Gosselin, 2002; Yamaguchi, Hirukawa, & Kanazawa, 1995). The similarity of these studies is that they focused on examining whether certain face parts are sufficient for gender categorization, while leaving the role of holistic processing unexplored. For example, Gosselin and colleagues asked participants to categorize the gender of a face presented behind a gray mask punctured by randomly located Gaussian apertures, so that only sparse face information was available. They found that the availability of eyes and eyebrows was significantly correlated with participants' correct gender decisions, indicating a diagnostic role of these facial features in gender categorization (e.g., Dupuis-Roy et al., 2009; Schyns et al.,

---

M. Zhao, mintao\_zhao@brown.edu

2002). Similarly, isolated face parts, particularly the eyes and brows, from averaged male and female faces have also been shown to contain gender information (Brown & Perrett, 1993; Yamaguchi et al., 1995). These findings suggest that gender processing may be part based, casting doubt on the generality of holistic face processing.

We employed multiple paradigms to infer the influence of holistic processing on face gender perception—specifically, the inversion, scrambling, and composite face effects. In each of these paradigms, the same physical features were available for different conditions; therefore, different performance between conditions cannot be attributed to the facial features displayed. If gender perception is holistic, we would expect an inversion effect (Experiment 1), a scrambling effect (Experiment 2), and a composite face effect (Experiment 3) in gender categorization similar to that in identity judgments. In contrast, if it is based on independent face parts, it should be relatively immune to the disruption of holistic face processing.

In all the experiments reported here, Chinese participants made gender judgments for both Chinese and Caucasian faces. Caucasian faces were included to ensure the generality of the results, since most previous studies of gender perception have tested Caucasian faces. The participants viewed faces that had undergone a transformation (inversion, scrambling, or placing one part of a face into a composite face) and were asked to make gender judgments for either the whole face (Experiments 1 and 2) or the top half of a face (Experiment 3).

## METHOD

### Participants

Thirty-two Chinese students from the University of Hong Kong participated in experiments for monetary reward. Experiments 1 and 2 each had 8 participants (3 males in each); Experiment 3 had 16 (7 males). Each participant was involved in only one experiment. All the participants reported normal or corrected-to-normal vision.

### Materials and Design

Photographs of unfamiliar Chinese ( $n = 80$ ; 40 males) and Caucasian ( $n = 80$ ; 40 males) students were converted to 8-bit grayscale and were masked by an oval shape, excluding hair and other peripheral details. The brightness and contrast of face images were roughly matched using Adobe Photoshop software. All these *intact* faces were standardized to an interpupillary distance of 60 pixels and were placed within a  $240 \times 315$  pixel black frame (Figure 1A). *Inverted* faces were created by vertically flipping the upright faces (Figure 1B). *Scrambled* faces were created by cutting each upright face into 10 components (two eyes, two eyebrows, two cheeks, nose, mouth, chin, and forehead) and rearranging these components into a new, unfacelike configuration (Figure 1C). Scrambled configurations were different from each other. The composite faces were created by first dividing each face into top and bottom segments by slicing them below the eyes and then interchanging the bottom segments with a randomly paired opposite gender face of the same race. The top and bottom segments were either *aligned* (Figure 1D) or *misaligned* (Figure 1E).

A  $2$  (race: Caucasian vs. Chinese)  $\times 2$  (version: upright vs. inverted in Experiment 1, intact vs. scrambled in Experiment 2, and aligned vs. misaligned in Experiment 3) within-participants design was used. The participants viewed 40 testing faces in each of the four conditions, and each testing face was displayed once for a participant. Across all participants in each experiment, testing faces were equally often presented in the two versions (e.g., upright vs. inverted in Experiment 1).

### Procedure

The participants were seated about 50 cm away from an eMac with a 17-in. CRT. Faces were presented one by one at the center of the screen (with a gray background) until a response was made. Trials were separated by a 1-sec blank screen. The participants were informed in advance about the versions of the faces that would be shown to them and the corresponding response keys for male and female faces. They were instructed to respond as accurately and as quickly as possible. The race and version of testing faces were blocked (e.g., Caucasian upright, Caucasian inverted, Chinese upright, and Chinese inverted in Experiment 1), with block orders counterbalanced across participants. The order of testing faces in each block was randomized. The participants received no feedback about their responses.

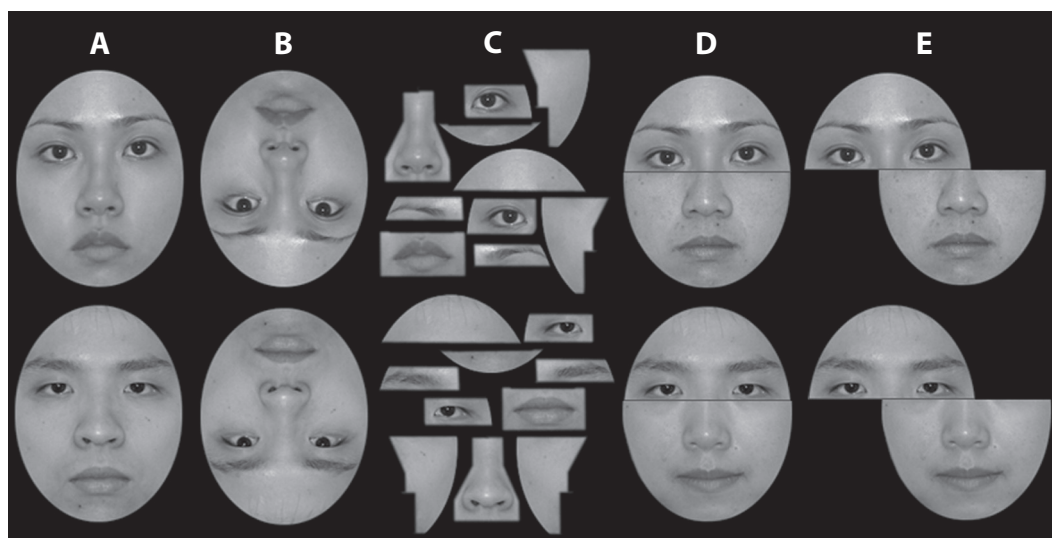


Figure 1. Examples of intact or upright (A), inverted (B), scrambled (C), aligned (D), and misaligned (E) faces of one Chinese female (top) and one Chinese male (bottom).

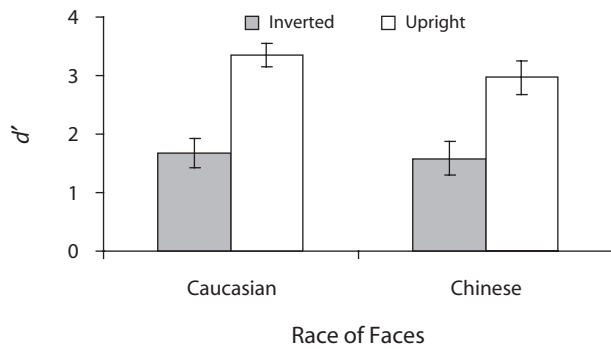


Figure 2. Response sensitivity ( $d'$ ) as a function of race and face orientation in Experiment 1. Error bars are standard errors of the means.

RESULTS AND DISCUSSION

The participants' performance was measured with  $d'$ , a signal detection measure of sensitivity; these data were then submitted to a repeated measures ANOVA, with race of face and stimuli version as within-participants factors. The participants' mean response latencies for correct responses (outliers more than 2.5 SDs above the mean were trimmed) were submitted to the same ANOVA to examine for possible speed-accuracy trade-offs.

Experiment 1: Upright Versus Inverted Faces

The participants in Experiment 1 judged the gender of upright and inverted faces. Response sensitivity was lower for inverted faces than for upright faces [ $F(1,7) = 605.72, MS_e = 0.07, p < .001, \eta_p^2 = .99$ ], showing an inversion effect (Figure 2). This inversion effect on gender perception was similar for Chinese and Caucasian faces, since no interaction between race and orientation was observed [ $F(1,7) = 2.47, p = .16, \eta_p^2 = .26$ ]; the participants showed equivalent overall performance for Chinese and Caucasian faces [ $F(1,7) = 1.47, p = .26, \eta_p^2 = .17$ ]. Because of concerns about the normality of the underlying population of  $d'$  scores, we checked individual participants' data to ensure that there was a consistent pattern; all the participants showed better sensitivity to upright than to inverted faces. Data from individual participants for all the experiments are presented in the Appendix.

Response latencies showed a similar inversion effect (Table 1, left): The participants spent more time in judging the gender of inverted than of upright faces [ $F(1,7) = 7.04, MS_e = 35,379, p < .05, \eta_p^2 = .50$ ]. This inversion effect seemed more prominent for Caucasian faces than for

Chinese faces, showing a marginally significant interaction between orientation and race [ $F(1,7) = 4.62, p = .07, \eta_p^2 = .40$ ]. Whether the test faces were Chinese or Caucasian had no effect on overall gender judgments ( $F < 1$ ).

Experiment 1 showed a clear inversion effect in face gender perception and an equivalent disruption effect due to inversion for Chinese and Caucasian faces. Faces were more difficult to categorize by gender when they were inverted than when they were upright. This result is consistent with similar findings when participants are asked to do identity judgments about a face (Rhodes, Hayward, & Winkler, 2006; Yin, 1969).

Experiment 2: Intact Versus Scrambled Faces

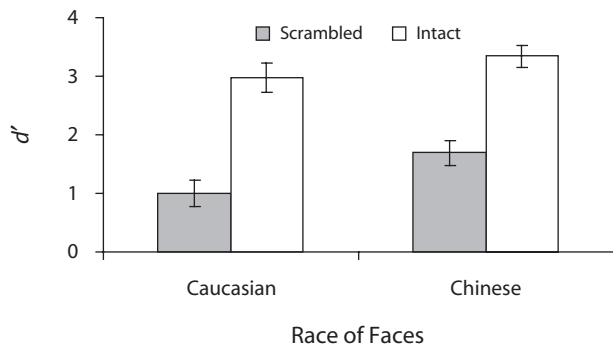
The participants in Experiment 2 judged the gender of intact and scrambled faces. Scrambling a face produced a decrement on gender perception similar to that produced by face inversion (Figure 3). The participants' performance in gender judgments dropped substantially for scrambled faces, as compared with intact faces [ $F(1,7) = 133.72, MS_e = 0.29, p < .001, \eta_p^2 = .95$ ]. The participants showed a trend of better performance in judging the gender of Chinese than of Caucasian faces [ $F(1,7) = 3.91, p = .09, \eta_p^2 = .36$ ]. This numerical difference between races was mainly due to judgments of scrambled faces [ $t(7) = 2.53, p < .05$ ], whereas judgments of intact faces were similar ( $t < 1$ ). However, the size of the scrambling decrement was similar for Chinese and Caucasian faces ( $F < 1$ ). All the participants showed the advantage for intact over scrambled faces.

Response latencies showed a similar scrambling effect (Table 1, middle). The participants were slower to judge the gender of scrambled than of intact faces [ $F(1,7) = 10.38, MS_e = 15,940, p < .02, \eta_p^2 = .60$ ]. The interaction between face version and race was significant [ $F(1,7) = 8.52, p = .02, \eta_p^2 = .55$ ], since the scrambling effect was larger for Caucasian faces than for Chinese faces. Response latencies for Chinese and Caucasian faces were similar ( $F < 1$ ).

Experiments 1 and 2 showed that face scrambling and face inversion have similar disruption effects on judgments of face gender, providing convergent evidence for the holistic nature of gender processing. These results demonstrated that face gender perception suffers dramatic impairments when holistic processing is eliminated or disrupted. The holistic nature of gender processing would be further supported if people found it difficult to ignore distracting holistic information in judging the gender of a face part. Experiment 3 demonstrated this mandatory nature of holistic gender processing with the composite face

Table 1 Mean Response Latencies (in Milliseconds; With Standard Deviations) As a Function of Testing Conditions in Experiments 1, 2, and 3

Test Face	Experiment 1				Experiment 2				Experiment 3			
	Inverted		Upright		Scrambled		Intact		Aligned		Misaligned	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Caucasian	1,209	458	844	215	1,996	1,060	708	122	1,202	374	1,134	402
Chinese	1,029	274	951	388	1,987	1,386	960	459	1,202	392	1,116	341



**Figure 3.** Response sensitivity ( $d'$ ) as a function of race and stimulus version in Experiment 2. Error bars are standard errors of the means.

paradigm, considered one of the most compelling demonstrations of holistic processing (Maurer et al., 2002).

### Experiment 3: Aligned Versus Misaligned Composite Faces

The participants in Experiment 3 were instructed to focus their attention on the top segment of aligned and misaligned faces and to judge the gender of these segments. Gender judgments for the top face segment showed more impairment when it was aligned with the bottom face segment of opposite gender than when they were misaligned [ $F(1,15) = 13.17$ ,  $MS_e = 0.18$ ,  $p < .005$ ,  $\eta_p^2 = .47$ ], demonstrating a composite face effect (Figure 4). Again, the disruption effect was similar for Chinese and Caucasian faces, since no interaction between race and version was observed ( $F < 1$ ). The overall performance showed no difference between Chinese and Caucasian faces ( $F < 1$ ). Thirteen of 16 participants showed better performance for misaligned than for aligned composites. Race and alignment of testing faces showed no effect on response latencies (Table 1, right). Judgments for misaligned faces were numerically faster than those for aligned faces [ $F(1,15) = 1.59$ ,  $p = .23$ ,  $\eta_p^2 = .10$ ; all other effects,  $F_s < 1$ ].

Experiment 3 demonstrated a composite face effect in gender perception. Face components (including eyes and eyebrows) within an upright face context cannot be well accessed for gender judgments without the influence of other face parts, indicating the mandatory holistic nature of gender perception.

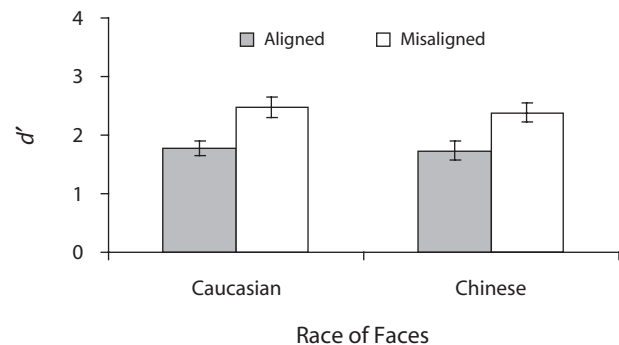
## GENERAL DISCUSSION

This study demonstrates the holistic nature of face gender perception, providing further evidence supporting the idea that faces are processed holistically. When holistic face processing was limited by inversion (Experiment 1), completely eliminated by scrambling (Experiment 2), or interfered with by an irrelevant face segment of the opposite gender (Experiment 3), gender judgments of a face were substantially disrupted. These results indicate that gender information about a face, like identity, is encoded holistically. When this holistic encoding is disrupted, people encounter difficulties in extracting a

face gender property, leading to a lower accuracy and a longer response time. Furthermore, people cannot ignore this holistic encoding even though they are instructed to selectively attend to the top part of a face. This is the first time, to our knowledge, that all these hallmark holistic-processing effects have been demonstrated along the visual-processing route to face gender categorization (Bruce & Young, 1986). Together with previous findings showing holistic-processing effects along visual analysis routes to judging face identity and expressions (e.g., Calder, Young, Keane, & Dean, 2000; Young et al., 1987), the present study indicates that holistic processing is a general mechanism of face perception that is recruited for both individualizing face identity and categorizing identity-independent aspects of a face.

The finding of holistic gender perception puts a fundamental constraint on interpreting roles of part-based visual cues in face gender classification. It is worth noting that the holistic nature of gender perception does not contradict the proposal that certain face parts play more dominant roles in gender decision than do others (e.g., Dupuis-Roy et al., 2009; Schyns et al., 2002) or the idea that part-based face information contributes to gender recognition (Brown & Perrett, 1993; Yamaguchi et al., 1995). In addition, this study does not rule out a role for other forms of componential face processing, such as that derived from principal component analysis (e.g., Valentin, Abdi, Edelman, & O'Toole, 1997). Consistent with studies of part-based identity processing (e.g., Hayward, Rhodes, & Schwaninger, 2008; Schwaninger et al., 2002), gender recognition based on face components (i.e., scrambled face parts) was also better than chance. However, the present study indicates that part-based processing is modulated by holistic processing of the face, even for gender perception. Those part-based visual cues work much better when they are situated in a normal upright face context (Experiments 1 and 2); furthermore, they cannot be selectively used for gender discrimination without the influence of other components within a face (Experiment 3).

The participants showed similar holistic-processing effects for Chinese and Caucasian faces, indicating the generality of holistic face gender perception. Nevertheless, this finding is inconsistent with studies of face identity, which



**Figure 4.** Response sensitivity ( $d'$ ) as a function of race and alignment in Experiment 3. Error bars are standard errors of the means.



have demonstrated stronger holistic-processing effects for faces of extensive contact experiences (e.g., Michel, Rossion, Han, Chung, & Caldara, 2006; Tanaka, Kiefer, & Bukach, 2004). Two possibilities might account for the discrepancy between holistic identity and gender processing. First, the holistic information used to encode gender and identity may be different even though both encodings are holistic. This independence has been demonstrated between holistic identity perception and holistic expression categorization (e.g., Calder et al., 2000; see also McKone, 2009; White, 2000). Second, contact experiences with faces may refine holistic processing only for subordinate-level individualization, but not for basic- or superordinate-level categorization that is based on characteristics shared by a group of examples. This assumption has been supported by Wong, Palmeri, and Gauthier's (2009) recent study, which showed that experiences for subordinate-level identification, rather than basic-level categorization, play a key role in developing holistic processing for nonface artificial objects.

Finally, there was little benefit of contact experience for our Chinese participants in judging the gender of Chinese faces, except for scrambled faces in Experiment 2, which is consistent with recent findings that contact experience with own-race faces helps with the analysis of scrambled faces (Hayward et al., 2008). A benefit for categorizing the gender of own-race normal upright faces has previously been shown by O'Toole, Peterson, and Deffenbacher (1996). Nevertheless, as was pointed out by Zhao and Bentin (2008), O'Toole et al.'s results might come from their specific procedures and heterogeneous Asian participants. Both Zhao and Bentin and the present study showed that contact experiences do not modulate gender classification of normal upright Chinese and Caucasian faces, consistent with the proposal that physiognomic properties that determine the gender of a face are universal, rather than race specific (Zhao & Bentin, 2008).

In summary, the present study demonstrated that face gender perception is holistic in nature. This finding provides evidence for the generality of holistic processing for different judgments of a face and the claim that faces are processed holistically (e.g., Farah et al., 1998). However, in contrast to face identity processing, holistic gender perception shows no effect of contact experience, echoing the idea that gender is defined with race-independent physiognomic properties (Zhao & Bentin, 2008) and that holistic information recruited for perceiving identity-independent aspects of a face may be different from that for analyzing identity (Calder et al., 2000).

**AUTHOR NOTE**

This research was supported by a grant from the Hong Kong Research Grants Council to W.G.H. (HKU4653/05H). We thank Bradley Gibson and three anonymous reviewers for their helpful comments on an earlier version of this article. Please address correspondence to M. Zhao, Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence, RI 02912 (e-mail: mintao\_zhao@brown.edu).

**REFERENCES**

BROWN, E., & PERRETT, D. I. (1993). What gives a face its gender? *Perception*, **22**, 829-840. doi:10.1068/p220829

BRUCE, V., & YOUNG, A. (1986). Understanding face recognition. *British Journal of Psychology*, **77**, 305-327.

CALDER, A. J., YOUNG, A. W., KEANE, J., & DEAN, M. (2000). Configurational information in facial expression perception. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 527-551. doi:10.1037/0096-1523.26.2.527

DUPUIS-ROY, N., FORTIN, I., FISET, D., & GOSSELIN, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, **9**(2:10), 1-8. doi:10.1167/9.2.10

FARAH, M. J., WILSON, K. D., DRAIN, M., & TANAKA, J. N. (1998). What is "special" about face perception? *Psychological Review*, **105**, 482-498. doi:10.1037/0033-295X.105.3.482

HAYWARD, W. G., RHODES, G., & SCHWANINGER, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition*, **106**, 1017-1027. doi:10.1016/j.cognition.2007.04.002

MAURER, D., LE GRAND, R., & MONDLOCH, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, **6**, 255-260. doi:10.1016/S1364-6613(02)01903-4

McKONE, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research*, **49**, 268-283. doi:10.1016/j.visres.2008.10.020

MICHEL, C., ROSSION, B., HAN, J., CHUNG, C.-S., & CALDARA, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, **17**, 608-615. doi:10.1111/j.1467-9280.2006.01752.x

O'TOOLE, A. J., PETERSON, J., & DEFFENBACHER, K. A. (1996). An "other-race effect" for categorizing faces by sex. *Perception*, **25**, 669-676. doi:10.1068/p250669

RHODES, G., HAYWARD, W. G., & WINKLER, C. (2006). Expert face coding: Configurational and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, **13**, 499-505.

SCHWANINGER, A., LOBMAIER, J. S., & COLLISHAW, S. M. (2002). Role of featural and configural information in familiar and unfamiliar face recognition. In H. H. Bülthoff, S.-W. Lee, T. A. Poggio, & C. Wallraven (Eds.), *Biologically motivated computer vision* (pp. 643-650). New York: Springer. doi:10.1007/3-540-36181-2\_64

SCHYNS, P. G., BONNAR, L., & GOSSELIN, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, **13**, 402-409. doi:10.1111/1467-9280.00472

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

TANAKA, J. W., KIEFER, M., & BUKACH, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, **93**, B1-B9. doi:10.1016/j.cognition.2003.09.011

VALENTIN, D., ABDI, H., EDELMAN, B., & O'TOOLE, A. J. (1997). Principal component and neural network analyses of face images: What can be generalized in gender classification? *Journal of Mathematical Psychology*, **41**, 398-412. doi:10.1006/jmps.1997.1186

WHITE, M. (2000). Parts and wholes in expression recognition. *Cognition & Emotion*, **14**, 39-60. doi:10.1080/026999300378987

WONG, A. C.-N., PALMERI, T. J., & GAUTHIER, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type? *Psychological Science*, **20**, 1108-1117. doi:10.1111/j.1467-9280.2009.02430.x

YAMAGUCHI, M. K., HIRUKAWA, T., & KANAZAWA, S. (1995). Judgment of gender through facial parts. *Perception*, **24**, 563-575. doi:10.1068/p240563

YIN, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, **81**, 141-145. doi:10.1037/h0027474

YOUNG, A. W., HELLAWELL, D., & HAY, D. C. (1987). Configurational information in face perception. *Perception*, **16**, 747-759. doi:10.1068/p160747

ZHAO, L., & BENTIN, S. (2008). Own- and other-race categorization of faces by race, gender, and age. *Psychonomic Bulletin & Review*, **15**, 1093-1099. doi:10.3758/PBR.15.6.1093

**NOTE**

1. Here, we use the term *holistic processing* to refer collectively to first-order relational, second-order relational, and holistic processing (Maurer et al., 2002).

**APPENDIX**  
**Response Sensitivity ( $d'$ ) for Individual Participants (P)**  
**in Experiments 1, 2, and 3**

P	Experiment 1		Experiment 2		Experiment 3	
	Upright	Inverted	Intact	Scrambled	Misaligned	Aligned
1	3.42	2.07	3.60	2.42	2.10	1.67
2	3.61	1.96	2.82	0.93	1.67	1.56
3	2.08	0.67	2.93	1.34	2.29*	2.38
4	2.41	0.54	2.90	1.22	2.81	1.90
5	3.76	2.42	3.14	1.44	1.34*	2.10
6	3.30	1.82	3.46	0.72	1.97	1.58
7	3.76	2.12	3.05	1.35	3.02	1.92
8	2.96	1.42	3.36	1.37	2.81	1.74
9					2.93	2.22
10					3.45	2.04
11					2.33	1.27
12					1.96*	2.32
13					2.48	0.65
14					2.66	0.81
15					3.27	2.14
16					1.87	1.75

\*Performance deviating from the overall pattern for each experiment.

(Manuscript received March 2, 2009;  
revision accepted for publication October 25, 2009.)