Perceptual Integration and the Composite Face Effect

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD-19-142.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	01-Nov-2019
Complete List of Authors:	Liu, Chang; Bournemouth University, Young, Andrew; University of York, Basra, Govina; Bournemouth University Ren, Naixin; Institute of Psychology, Chinese Academy of Sciences, Chen, Wenfeng; Renmin University of China, Department of Psychology
Keywords:	composite face, holistic processing, face perception, face gender, face race

DOI: 10.1177/1747021819899531

REVISION: 29.10.19.

Perceptual Integration and the Composite Face Effect

Chang Hong Liu

Bournemouth University

Andrew W. Young

University of York

Govina Basra

Bournemouth University

Naixin Ren

Chinese Academy of Sciences

Wenfeng Chen

Renmin University of China

Author Notes

Chang Hong Liu and Govina Basra, Department of Psychology, Bournemouth University. Fern Barrow, Poole, Dorset, BH12 5BB, UK.

Andrew W. Young, Department of Psychology, University of York, York YO10 5DD, UK.

Naixin Ren, State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences.

The research was supported by grants from the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China (18XNLG10) and the National Natural Science Foundation of China (31371031, 61632004) .

Correspondence concerning this article should be addressed to Wenfeng Chen, Department of Psychology, Renmin University of China, No. 59 Zhongguancun Street, Haidian District, Beijing 100872, P.R. China

Keywords: composite effect, holistic processing, face perception, face gender, face race

DOI: 10.1177/1747021819899531

Abstract

The composite face paradigm is widely used to investigate holistic perception of faces. In the paradigm, parts from different faces (usually the top and bottom halves) are recombined. The principal criterion for holistic perception is that responses involving the component parts of composites in which the parts are aligned into a face-like configuration are disrupted compared to the same parts in a misaligned (not face-like) format. This is often taken as evidence that seeing a whole face in the aligned condition interferes with perceiving its separate parts, but the extent to which the effect is perceptually driven remains unclear. We used salient perceptual categories of gender (male or female) and race (Asian or Caucasian appearance) to create composite stimuli from parts of faces that varied orthogonally on these characteristics. In Experiment 1, participants categorized the gender of the parts of aligned composite and misaligned images created from parts with the same (congruent) or different (incongruent) gender and the same (congruent) or different (incongruent) race. In Experiment 2 the same stimuli were used but the task changed to categorizing race. In both experiments there was a strong influence of the task-relevant manipulation on the composite effect, with slower responses to aligned stimuli with incongruent gender in Experiment 1 and incongruent race in Experiment 2. In contrast, the task-irrelevant variable (race in Experiment 1, gender in Experiment 2) did not exert much influence on the composite effect in either experiment. These findings show that whilst holistic integration of salient visual properties makes a strong contribution to the composite face effect, it clearly also involves targeted processing of an attended visual characteristic.

DOI: 10.1177/1747021819899531

Introduction

Variants of the composite face paradigm (Young, Hellawell & Hay, 1987; Hole, 1994) now form the most widely-used technique for investigating holistic perception of faces (McKone & Robbins, 2011; Murphy, Gray, & Cook, 2016; Rossion, 2013; Tanaka & Gordon, 2011). Typically, the top part of one face is aligned with the bottom part of another face to create a face-like composite image or the same combination of face parts is presented in a misaligned format that does not look face-like. Participants are then asked to make decisions about the face parts themselves. The highly replicable pattern of findings is one of slower and/or less accurate responses in the aligned than misaligned format condition, and this is often taken to indicate that holistic perception of the face-like aligned stimulus interferes with perceiving its constituent parts. In the misaligned condition there is thought to be no interference from holistic face perception because the stimulus does not look face-like. Hence the difference between responses to the same face parts presented in aligned and misaligned formats is taken to form an index of holistic perception.

Although interpretations of the composite face effect in terms of holistic perception still dominate the published research literature, and for some years it was considered to offer a 'gold standard' measure of holistic perception (Avidan, Tanzer & Behrmann, 2011; McKone & Robbins, 2011), it has become clear that the composite effect is to some degree also susceptible to other sources of influence. In essence, the paradigm pits holistic perception of the aligned composite face against featural perception of its constituent parts and in consequence involves not only holistic perception but also ability to selectively attend and respond to the face parts specified in the experimental task. The review by Murphy et al. (2016) summarizes a number of studies implicating the contribution of such factors alongside holistic perception *per se*.

The importance of issues surrounding how to measure holistic perception using the composite effect can be clearly seen in the controversy concerning measures based on matching tasks. The composite face paradigm as originally introduced by Young et al. (1987) used a single aligned or misaligned composite stimulus created from photographs of familiar faces and simply asked participants to identify the top or bottom part. However the restrictions inherent in this familiar face recognition task led Hole (1994) to introduce a

DOI: 10.1177/1747021819899531

useful variant in the form of a matching paradigm in which pairs of aligned or misaligned composite stimuli are created from unfamiliar faces and the participant is asked whether their top or the bottom parts are the same. In Hole's (1994) study the pairs of composite images were presented simultaneously, but variants of this matching task involving successive stimulus presentation with a short inter-stimulus interval have since become the most widely used and led to a number of important findings (Rossion, 2013). Despite this widespread use of matching tasks, though, differences of opinion have arisen concerning which combinations of top and bottom parts are essential to a valid measure (the 'partial' vs. 'complete' designs; see Richler & Gauthier, 2014) and related measurement issues such as how to make use of 'same' and 'different' responses (Richler & Gauthier, 2013; Rossion, 2013). What has become clear from this controversy is that different measures can lead to different outcomes; they do not seem to offer direct indices of holistic perception uncontaminated by other influences.

A further complication is that the concept of holistic perception can itself be interpreted in different ways (Cheng, McCarthy, Wang, Palmeri & Little, 2018; Maurer, Le Grand & Mondloch, 2002; Richler, Palmeri & Gauthier, 2012). Often, the composite face effect is taken to indicate the existence of an integral overall perceptual representation of the face that does not explicitly represent its constituent parts. However, much of the evidence is consistent with an interpretation in terms of a combination of serial and parallel processing of facial features themselves (Cheng et al., 2018).

As Richler and Gauthier (2014, p. 1281) emphasise, the concept of holistic processing has become a cornerstone of face recognition research. Establishing how different factors contribute to the composite face effect is therefore important both to understanding holistic perception and to determining how best to try to measure it. Here, we follow a variant of Young et al.'s (1987) original procedure introduced by Calder, Young, Keane and Dean (2000) which involves creating combinations of parts that are task-relevant or task-irrelevant and investigating how these influence the composite face effect. Reverting to this procedure derived directly from Young et al. (1987) eliminates the influence of some of the factors that contribute to performance in matching tasks. In particular, the use of a single composite stimulus on each experimental trials removes the short-term memory component of sequential matching paradigms and it also removes the inter-stimulus comparison component

DOI: 10.1177/1747021819899531

required for deciding whether parts are 'same' or 'different' in simultaneous or successive matching. These task components cannot then influence the obtained measures.

Our approach is derived from Calder et al. (2000), who created composite images in which the top and bottom halves could show the same or a different emotional expression and the same or a different face identity. They found a composite face expression effect in which participants were slower to identify the expression in either half of these composite images relative to a misaligned control condition. Likewise, they found a composite face identity effect in which participants were slower to identify the identity in either half of these composite images relative to a misaligned control condition. Importantly, however, when Calder et al. (2000) investigated the relation between the composite expression effect and the composite effect for face identity, they found that these were independent of one another. In other words, the composite face expression effect was not affected by differences in the identities of the constituent parts, and, conversely, when participants were asked to recognize face identity, the composite face identity effect was no longer affected by differences in the expressions. Hence the composite effect was determined by task-relevant differences between the parts (identity differences in the identity task and expression differences in the expression task) and unaffected by task-irrelevant differences between the parts (expression differences in the identity task and identity differences in the expression task).

A development of this approach was made by Chen, Ren, Young and Liu (2018), who created composites from parts that varied in gender (male or female) and occupation (movie star or athlete). Their findings, however, were strikingly different from the pattern noted by Calder et al. (2000). Instead, Chen et al. (2018) found that gender differences between the top and bottom parts of the stimuli led to a face composite effect (i.e. slower responses to aligned than misaligned stimuli) regardless of whether the task involved judging the gender, occupation or identity of the parts. In contrast, differences in occupation between the parts exerted only a relatively limited influence on the composite effect even when the task was to judge occupation. The influences of gender and occupation differences on the composite effect were therefore asymmetric, with gender differences having a more pervasive effect than differences in occupation.

Chen et al. (2018) interpreted their findings in terms of Bruce and Young's (1986)

DOI: 10.1177/1747021819899531

distinction between visually-derived and identity-specific semantic information. The concept of visually-derived semantic information refers to perceptual properties (of which gender is a prime example) that differ more or less consistently across social categories (in this case male and female faces) and can therefore be seen in unfamiliar or familiar faces without needing to recognize the face's identity as such. Identity-specific semantic information refers to characteristics (such as occupation) that can only be accurately determined by accessing previously stored knowledge through recognizing the identity of a familiar face. From this perspective the pervasive influence of the visually-derived property of gender congruence on the composite face effect in Chen et al.'s (2018) experiments shows how interference originating in a relatively early perceptual locus can be a dominant factor in the composite effect. Since Chen et al. (2018) chose faces that minimized perceptual differences across occupations, the inconsistency of the identity-specific occupation congruence effects points to a more minor role for non-perceptual decision and response factors.

Looked at in this way, Calder et al.'s (2000) findings of independent composite effects for face identity and expression may well derive from the fact that both have strong perceptual correlates. However, an alternative possibility is that expression and identity are often considered to represent aspects of face perception that require distinct types of perceptual representation (cf. Bruce and Young, 1986). Although it is now clear that there is some overlap in the perceptual bases of identity and expression (Calder et al., 2001; Calder & Young, 2005) this overlap is clearly partial and there are a number of fundamental differences between the functional demands of identity and expression processing (Young, 2018).

We therefore decided to test Chen et al.'s (2018) interpretation by creating composite stimuli from orthogonally varied parts of faces of both different gender (male or female) and different race (Asian or Caucasian appearance). Although both gender and race are complex social constructs and modern approaches to biogenetic diversity focus more on geographically defined subpopulations than the idea of race *per se*, gender and race also form salient perceptual and social categories that are easily seen in unfamiliar faces (Bruce & Young, 2012; Freeman & Ambady, 2011) and can be computed from image properties (Kramer, Young, Day & Burton, 2017). In Bruce and Young's (1986) terms gender and race

DOI: 10.1177/1747021819899531

are quintessential forms of visually-derived semantic information.

In Experiment 1, then, we asked participants to classify the gender of the top or the bottom parts of aligned composite and misaligned images created from parts that were of the same (congruent) or different (incongruent) gender and the same (congruent) or different (incongruent) race. Our interest was in the extent to which the composite effect would be influenced by the task-relevant variable of gender congruence and by the task-irrelevant variable of race congruence. In Experiment 2 we employed exactly the same stimuli but changed the task to one of classifying the race of the top or the bottom parts of the aligned composite and misaligned images, so that gender congruence was now task-irrelevant and race congruence became task-relevant.

Experiment 1

In a study of holistic representation of face gender, Baudouin and Humphreys (2006) created gender-congruent (male top part and male bottom part, or female top part and female bottom part) and gender-incongruent stimuli (male and female parts) from aligned and misaligned top and bottom half face parts. When participants were asked to categorize the gender of the face parts, Baudouin and Humphreys (2006) found that the difference between responses to aligned and misaligned stimuli was greater for the gender-incongruent than for the gender-congruent stimuli. This was taken to reflect a composite face effect in which the perception of a new whole face of indeterminate gender for aligned stimuli in the gender-incongruent condition interfered with classifying its constituent parts.

Chen et al. (2018) adapted this paradigm to create a new variant involving composite faces drawn from orthogonal social categories of gender and occupation. Their aim was to assess the contributions from visually-derived semantic properties (gender) or identity-specific semantic properties (occupation) to composite interference and to measure whether the task-irrelevant category could moderate the composite effect. They found evidence of interference from gender when it was task-irrelevant, but little effect from occupation when it was task-irrelevant.

In the present experiment, we investigated whether the limited effect of occupation found by Chen et al. (2018) was due to the lack of consistent visual differences associated

DOI: 10.1177/1747021819899531

with this semantic information. We employed the same gender categorization task but replaced occupation with face race as the task-irrelevant orthogonal category. Like gender, face race involves salient visual properties.

As well as seeking to extend Chen et al.'s (2018) approach by investigating the effects of task-relevant gender congruence and task-irrelevant race congruence in a gender categorization task, we tested whether the results could be generalized to two different ethnic groups. Chen et al. (2018) had only tested Chinese participants, so we sought to check whether a comparable pattern of findings would be made for British participants. We considered this to be important because there are substantial cultural differences between East Asia and the West (Hofstede, 1980; Nisbett, Peng, Choi & Norenzayan, 2001; Oyserman, Coon, & Kemmelmeier, 2002) that have been found in previous studies to affect face perception and resulting social judgments (Blais et al., 2008; Chen, Jing, & Lee, 2012; Jack et al., 2009, 2012; Wheeler & Kim, 1997).

Method

Participants. These were 20 Chinese university students (Mean age 23.7 years, range 19-27 years, 10 females) and 21 British university students (Mean age 19.6 years, range 18-22 years, 20 females). All had normal or corrected-to-normal vision. All participants in the Chinese group were of Han Chinese origin (the main ethnic group in China), and all in the British group were of Caucasian origin. The study was conducted in accordance with the APA's guidelines on the treatment of human participants and was approved by the appropriate local ethics committees.

Materials and design. Eight photographs of unfamiliar faces, four Chinese and four Caucasian, with equal numbers of males and females of each race, were used to create the stimuli. All faces had a neutral expression. The photographs were converted to grayscale images and were edited in Photoshop to remove hair and external features. The background was then replaced with a uniform neutral grey. The overall image size was normalized to 120 × 164 pixels, which subtended 3.61 × 4.84°of visual angle under the experimental conditions.

Figure 1 shows examples of stimuli in each condition. Each original face image was divided into top and bottom halves along a horizontal line through the bridge of the nose. The top half of one face was then aligned with the bottom half of a different face to produce

DOI: 10.1177/1747021819899531

dace-like aligned composite images. Following a recommendation by Rossion (2013), a small gap (0.1°) was kept between the two halves, so that the line separating the top from the bottom part of the composite faces would always be clear to participants. Corresponding misaligned stimuli were then created from the same combinations of face halves by offsetting the top and bottom face parts, such that the center of the bottom half was shifted to line up vertically with one edge of the top half. The resulting misaligned images have a much less face-like appearance than the aligned composites.

The design involved comparing reaction times and error rates to classify the top or bottom parts of aligned and misaligned stimuli. These top and bottom parts were themselves of congruent gender (both parts male or both parts female) or incongruent gender (one part male and one part female) or of congruent race (both parts Asian or both parts Caucasian) or incongruent race (one part Asian and one part Caucasian). A fully crossed design was used in which all possible combinations of congruent and incongruent gender and race were represented (see Figure 1).

Each face was used once as a top half and once as a bottom half in each combination of gender and race congruence. The top and bottom parts of all stimuli were always from two different identities. As Figure 1 shows, the combinations of congruent or incongruent gender and congruent or incongruent race resulted in four types of pairing: 1) Congruent gender, congruent race (e.g., a top part of a Chinese male paired with a bottom part of another Chinese male); 2) Congruent gender, incongruent race (e.g., a top part of a Chinese male paired with a bottom part of a Caucasian male); 3) Incongruent gender, congruent race (e.g., a top part of a Chinese male paired with a bottom part of a Chinese female); 4) Incongruent gender, incongruent race (e.g., a top part of a Chinese male paired with a bottom part of a Caucasian female). This led to a total of 32 aligned composite pairings (Figure 1A) and 64 misaligned composites resulting from a leftwards or rightwards shift of the parts (Figure 1B). Only half of the misaligned composites are shown in Figure 1B.

The main experimental trials required participants to judge the gender of either the top half or the bottom half of these aligned composite and misaligned images. When the top half was to be judged, it was always presented centered across the vertical midline of the screen (for both aligned and misaligned stimuli), whereas the bottom half of the misaligned stimuli

DOI: 10.1177/1747021819899531

was either shifted to the left or to the right, with the two directions of shift counterbalanced. When the bottom half was to be judged, it was always presented centered across the vertical midline of the screen, whereas the top half of misaligned stimuli was either shifted to the left or to the right. In this way, the location of the part of the stimuli to be judged was always the same for the top half judgments and always the same for the bottom half judgments. Moreover, the specific combinations of top and bottom half image pairings were the same for the aligned composite and misaligned conditions, so that any differences between responses to aligned and misaligned stimuli must be directly attributable to this variable.

Procedure. Participants were asked to categorize the gender of the top halves of the stimuli or the bottom halves, in separate blocks of trials.

The general procedure followed that used by Chen et al. (2018). In an initial block of trials, participants judged whether the top or bottom half of a face was a male or a female when it was presented in isolation. This was done to ensure that participants were familiar with the task and able correctly to categorize the stimuli prior to the experimental trials themselves.

In the main blocks of experimental trials, participants categorized as male or female the top or bottom halves of the aligned and misaligned stimuli. Each participant was asked to judge the gender of the top or the bottom half of the face as quickly and accurately as possible by pressing one of two designated keys on the keyboard. The keys associated with male and female responses were counterbalanced across participants. As already noted, two separate blocks of trials were used, where participants either judged the top or the bottom half of the face stimuli. The order of these top and bottom blocks was counterbalanced across participants. Prior to each block, an instruction screen informed the participants whether they were expected to judge the top half or the bottom half of the face.

All trials in the two blocks followed the same procedure, whereby a 500ms fixation screen was followed by the stimulus to be categorized. Each stimulus was presented until a response was made. There was a 1-s inter-trial interval following each response. As noted above, each block was itself divided into two sessions. The first session showed isolated halves of faces, and was intended as a check that participants could correctly classify the top and bottom part of each face as male or female. The second session was the main

DOI: 10.1177/1747021819899531

experimental trials involving aligned composite or misaligned face parts.

Isolated halves: Only the top or the bottom half of a face was presented in this session, depending on whether the block required judging the gender of the top or the bottom half. Each stimulus was repeated four times to give a total of 32 trials in a random order, with the constraint that no image was repeated twice in a row.

Combined halves: Aligned composite and misaligned images created from the top and bottom halves of different faces were presented in this session. In each trial, a composite image was shown individually in the center of the screen. The 32 aligned composites were presented twice, and the 64 misaligned composites were presented once. This resulted in a total of 128 trials. The order of trials was otherwise random with the following constraints; the same face image was not allowed to appear repeatedly in two consecutive trials, and the same correct response key was only allowed to repeat up to five times in a row. The participant's task was to categorize the gender of the top part of each stimulus or the bottom part of each stimulus in the separate blocks of trials. Participants were asked to do this as quickly and accurately as possible, whilst ignoring the irrelevant part (i.e. ignoring the bottom part when asked to classify the top part, and vice versa). To familiarize participants with the task, this session was itself preceded by 16 practice trials.

In sum, for the main experimental trials we employed within-participant factors of Gender Congruence (congruent or incongruent male and female face parts), Race Congruence (congruent or incongruent Chinese and Caucasian face parts), Alignment (aligned or misaligned stimulus arrangement) and Position to be categorized (top-half or bottom-half). Participant Race (Chinese or Caucasian ethnic background) was a between-participant factor. The task was always gender categorisation.

Results

Isolated halves. First, we consider performance with the part faces presented in isolation; this was used to establish that the main task of classifying face parts could be accurately performed and to ensure that participants were familiar with the task before the main experimental trials. The accuracies of judging the gender of a face based on the top or bottom half of the face alone were 97% (*SD* = 4) and 96% (*SD* = 6), respectively. From this we conclude that the top and bottom halves of each of the images used to create the experimental

DOI: 10.1177/1747021819899531

stimuli contained cues that allowed them to be accurately classified as male or as female. However, classifying the top half was more accurate, $t(40) = 2.20$, $p = .034$, Cohen's $D = 0.43$. Correct responses for the isolated top half ($M = 738ms$, $SD = 181$) were noted to be faster than for the isolated bottom half ($M = 813ms$, $SD = 162$) of the faces, $t(40) = 2.77$, $p = .008$, Cohen's $D = 0.44$.

Combined halves. The task was intended to allow reasonably high accuracy, so that the most important findings involve reaction times. In line with criteria used in our previous study (Chen et al., 2018), reaction times from trials with incorrect responses and outliers of more than 2.5 SDs were excluded in the calculation of mean RTs in each condition; Omnibus test of multivariate normality of the data showed that the RTs were normally distributed ($p = .90$). A five-way repeated-measures analysis of variance (ANOVA) was used to examine the effects of stimulus Alignment (aligned or misaligned stimuli), Gender Congruence (congruent or incongruent combinations of male and female images), Race Congruence (congruent or incongruent combinations of Asian and Caucasian images), Participant Race (Chinese or British), and task Position (responses to top or bottom halves).

The ANOVA is summarized in Table 1. As the main interest of the experiment was in the composite effect, which is indexed by a difference between aligned and misaligned versions of the stimuli, we begin by reporting findings that involve the Alignment variable. There was a main effect of Alignment ($\eta^2_p = .80$ - see Table 1), representing the classic composite effect, which was qualified by an interaction with Gender Congruence ($\eta^2_p = .29$). Figure 2 illustrates this interaction. Simple effect analyses of this Gender Congruence x Alignment interaction revealed a significant effect of Alignment for incongruent-gender pairings, $F(1, 39) = 96.41$, $p < .001$, $\eta^2_p = .71$, where responses for the aligned stimuli were slower than for the misaligned. There was also a similar but a smaller effect of Alignment for congruent-gender pairings than for incongruent-gender pairings, $F(1, 39) = 18.94$, $p < .001$, $\eta^2_p = .33$. Another way to analyze the interaction is to look at the effect of Gender Congruence on aligned vs misaligned halves. This showed a significant effect of Gender Congruence when the two halves were aligned, where responses for incongruent gender were slower than for congruent gender, $F(1, 39) = 30.32$, $p < .001$, $\eta^2_p = .44$. In contrast, there was no effect of Gender Congruence for misaligned halves, $F(1, 39) = 3.37$, $p = .074$, $\eta^2_p = .08$.

DOI: 10.1177/1747021819899531

There was also an interaction between Alignment and Position (η^2_p = .17), which is illustrated in Figure 3. Although effects of Alignment were significant for both top- and bottom-halves, F's(1, 39) ≥ 20.58, p's < .001, η^2_p's ≥ .35, the Alignment effect for bottom-halves was larger.

Main effects of Gender Congruence (η^2_p = .44) and Position (η^2_p = .33) were also noted, but these are subsumed under the interactions presented in Figures 2 and 3. No other effects reached significance, though we should note a very marginal 4-way interaction of Gender Congruence, Race Congruence, Participant Race and Alignment (p = .097) with a small effect size (η^2_p = .07).

Although response times formed our principal focus of interest, errors were also analyzed as a check for any speed-accuracy trade-offs. The overall error rate was 14%. A five-way repeated-measures analysis of variance (ANOVA) of the arcsine-transformed accuracy data revealed an interaction of Gender Congruence × Alignment, $F(1, 39) = 34.37$, $p < .001$, $\eta^2_p = .47$. When the top and bottom parts of the stimuli consisted of different gender, response accuracies for misaligned stimuli (87%) were higher than for aligned stimuli (81%), $F(1, 39) = 40.42$, $p < .001$, $\eta^2_p = .51$. In contrast, when the two parts consisted of the same gender, the accuracies for aligned and misaligned stimuli were both 89%, with no effect of Alignment, $F(1, 39) = 0.03$, $p = .856$, $\eta^2_p < .01$. This interaction is therefore comparable in form to that found for reaction times, indicating that participants did not trade speed for accuracy.

The accuracy data ANOVA also revealed significant main effects of Alignment, $F(1, 39) = 21.55$, $p < .001$, $\eta^2_p = .36$, Gender Congruence, $F(1, 39) = 35.71$, $p < .001$, $\eta^2_p = .48$, Race Congruence, $F(1, 39) = 4.39$, $p = .043$, $\eta^2_p = .10$, and Position, $F(1, 39) = 33.40$, $p < .001$, $\eta^2_p = .46$, that were all in line with the findings from response times. However, there was also a significant two-way interaction between Gender Congruence and Race Congruence, $F(1, 39) = 14.99$, $p < .001$, $\eta^2_p = .28$. Simple effect analyses of this interaction of Gender congruence × Race congruence found that when two halves of a stimulus consisted of the same face race, the accuracy for congruent gender pairs (90%) was higher than for incongruent gender pairs (84%), $F(1, 39) = 24.30$, $p < .001$, $\eta^2_p = .38$. In contrast, no effect of this Gender Congruence was found (the means for congruent and incongruent gender were 88% and 85%, respectively)

DOI: 10.1177/1747021819899531

when the two halves of a stimulus consisted of different face races, $F(1, 39) = 2.02$, $p = .163$. No other main effects or interactions reached significance, $F's(1, 39) < 3.75$, $p's > .05$.

Discussion

The aim of Experiment 1 was to examine how the task-relevant visual category of gender and the task-irrelevant visual category of race would influence the composite effect. As expected, for the task-relevant gender category, we found a clear composite effect, where judgments about the gender of the top or bottom half of a face were slower when the two halves were aligned than misaligned. However, and as previously noted both by Baudouin and Humphreys (2006) and by Chen et al. (2018), this effect was most clearly evident when the gender of the two halves was incongruent (the Gender Congruence x Alignment interaction).

We also found that the alignment effect was greater when the task was to judge the bottom part of the face. This is consistent with the findings of Chen et al. (2018), where a similar pattern of results was reported, though the overall response times in the present study were longer. These longer response times are likely due to a difference in the type of face stimuli employed in the two studies. Chen et al. (2018) used familiar faces, whereas Experiment 1 used unfamiliar faces. Our current findings confirm that although the overall gender judgment may be slower, the gender congruence effect extends also to unfamiliar faces.

In contrast to task-relevant information, the task-irrelevant race congruence variable did not influence the composite face effect; that is, the task-irrelevant characteristic had little influence on participants' judgements. This suggests that participants were able to focus on the task-relevant holistic representation of face gender, while ignoring the task-irrelevant holistic representation of face race. Also, it shows that the gender congruence effect was equally created by both race-congruent and race-incongruent pairings. This extends the findings of Chen et al (2018), who demonstrated the gender congruence effect only using race-congruent stimuli.

Because Chen et al. (2018) had only tested participants of Chinese ethnicity, we also tested British participants in Experiment 1, in case there might be any substantial cultural differences. None were evident. Note though that we did not test the other-race effect, which

DOI: 10.1177/1747021819899531

although well-established in many other studies (Hayward, Crookes & Rhodes, 2013; Michel et al., 2006; Rossion & Michel, 2011) was not part of our focus of interest here. To test for other-race effects it would have been necessary to analyse the data in terms of Stimulus Race as well as Participant Race. This would have restricted the own-race condition to that half of the congruent race set in which the race of the face parts was the same as the race of the participant and therefore complicated the analyses and led to a substantial loss of power.

Chen et al. (2018) pointed out that the gender congruence effect is likely due to the incongruent gender pairings appearing more ambiguous in gender than the congruent gender pairings when the top and bottom halves are aligned. It was noteworthy that any comparable racial ambiguity created by the incongruent race pairs had minimal impact on the processing speed in this experiment, but of course the race dimension was irrelevant to the gender categorization task. We therefore sought in Experiment 2 to determine whether race incongruence might have a similar impact to gender incongruence when it was made task-relevant.

Experiment 2

In Experiment 2, we asked participants to categorize the race of the top or bottom halves of aligned or misaligned images, using the same the stimuli as in Experiment 1. In this way we made the face race now task-relevant, while the face gender became task-irrelevant. This allowed us to further investigate the effect of task relevance on holistic processing in face composites. If task relevance determines the influence of a salient perceptual category in the composite face paradigm, we would expect a reversal of the effects we found in Experiment 1, leading to a modulation of the composite face effect by the task-relevant race congruence, accompanied by a diminished influence of the task-irrelevant gender congruence.

Method

Participants. There were 20 Chinese university students (mean age 22.8 years, range 19-27 years, 10 females) and 27 British university students (Mean age 19.9 years, range 18-27 years, 25 females) who participated in the experiment. All had normal or corrected-to-normal vision. None had taken part in Experiment 1.

DOI: 10.1177/1747021819899531

Materials and design. The stimuli and experimental conditions were identical to those created for Experiment 1.

Procedure. The procedure was also identical to Experiment 1, except that the gender categorization task used for Experiment 1 was changed to a race categorization task in Experiment 2. Participants were required to classify the top or bottom part of each image in terms of the face race (Asian or Caucasian).

Results

Isolated halves. The mean accuracies for categorizing face race from the top and bottom halves were 98% (*SD* = 4) and 93% (*SD* = 10), respectively. Hence it is clear that the top and bottom halves of each of the images used to create the experimental stimuli contained cues that allowed them to be accurately classified as being of Asian or Caucasian appearance. However, the top halves (98%) were recognized more accurately than the bottom halves (93%), $t(46) = 3.01$, $p = .004$, Cohen's $D = 0.53$. The mean correct reaction time was also faster for the for top halves ($M = 677ms$, $SD = 194$) than for the bottom halves ($M = 887ms$, $SD = 318$), $t(46)= 4.37$, $p < .001$, Cohen's $D = 0.79$.

Combined halves. As for Experiment 1, our main analysis involves the reaction time data. Again, reaction times from trials with incorrect responses and outliers of more than 2.5 SDs were excluded, and Omnibus test of multivariate normality of the data showed that the RTs were normally distributed ($p = .99$). Table 2 shows the details of the five-factor repeated-measures ANOVA of correct RTs.

Again, we mainly focused on the effects involving the Alignment factor as an index of the composite effect. The ANOVA results showed a main effect of Alignment ($\eta^2_p = .57$) and significant interactions between Race Congruence and Alignment ($\eta^2_p = .30$) and between Alignment and Position ($\eta^2_p = .69$). These effects were however subsumed under a three-way interaction between the task-relevant factor Race Congruence, Alignment and Position ($\eta^2_p = .17$) shown in Figure 4. Simple effects analyses of this interaction showed that the Race Congruence × Alignment interaction was absent when the top half was judged, $F(1, 45) = .60$, $p = .44$, $\eta^2_p = .01$, but present when the bottom half was judged, $F(1, 45) = 119.71$, $p < .001$, $\eta^2_p = .73$.

The interaction between Alignment and the task-irrelevant Gender Congruence variable

DOI: 10.1177/1747021819899531

only reached a borderline level of statistical significance with a small effect size, p = .092, η^2_p = .06. Simple effects analyses showed that Alignment effects in both congruent- and incongruent-gender conditions were significant, F's(1, 45) ≥ 29.79, p's < .001, $\eta^2_p \geq .40$; the marginal interaction appeared to be due to a slightly stronger alignment effect for the incongruent gender pairings relative to the congruent gender pairings.

As Table 2 shows, there were also some significant results that did not involve the Alignment variable. These effects are of less interest, but we provide a brief summary here. One of these was the two-way interaction between the task-relevant Race Congruence and task-irrelevant Gender Congruence. However, this was qualified by a four-way interaction of Participant Race, Race Congruence, Gender Congruence, and Position. Hence our further analyses were focused on this four-way interaction. We conducted simple effect analyses for the two positions. These showed a significant three-way interaction of Participant Race, Race Congruence and Gender Congruence for the Top condition, $F(1, 45) = 13.54$, $p = .001$, $\eta^2_p = .23$, but not for the Bottom condition, $F(1, 45) < 0.01$, $p = .980$, $\eta^2_p < .01$. Analysis of this three-way interaction showed a Race Congruence × Gender Congruence interaction in Chinese participants, $F(1, 45) = 18.08$, $p < .001$, $\eta^2_p = .29$, but not in British participants, $F(1, 45) = 0.49$, $p = .486$, $\eta^2_p = .01$.

We next analyzed the cause of this Race Congruence × Gender Congruence interaction in the Chinese group. The results showed that Chinese participants classified the race of the Top half more quickly when both the race and gender of a composite face were congruent (624ms) than when the congruent factor was only gender (652ms) or race (658ms), F's(1, 45) ≥ 9.38, p's ≤ .004, $\eta^2_p \geq .17$. When both race and gender were incongruent, the Top half was classified more quickly by Chinese participants (636ms) than when the incongruent factor was only race (652ms) or gender (658ms), F's(1, 45) = 5.32, p's ≥ .026, $\eta^2_p \geq .10$. We note thought that this interaction effect was found only for Chinese participants and only when they classified the top of the composite faces.

The overall error rate in this session was 6%. A five-way ANOVA of the arcsine transformed accuracy data found significant main effects of Alignment, $F(1, 45) = 11.45$, $p = .001$, $\eta^2_p = .20$, Race Congruence, $F(1, 45) = 48.52$, $p < .001$, $\eta^2_p = .52$, and Position, $F(1, 45) = 55.33$, $p < .001$, $\eta^2_p = .55$, which were all in line with the response time results. There

DOI: 10.1177/1747021819899531

were also significant two-way interactions between Race Congruence and Position, and between Race Congruence and Alignment, but both were implicated in a three-way interaction between Race Congruence, Position, and Alignment, $F(1, 45) = 25.33$, $p < .001$, $\eta^2_p = .36$.

To analyze this three-way interaction, we looked at the Top and Bottom results separately. When participants classified the race of the Bottom half, the Race Congruence × Alignment interaction was significant, $F(1, 45) = 25.01$, $p < .001$, $\eta^2_p = .36$. This was due to the typical composite face effect, where race categorization for aligned halves was less accurate (83%) than for misaligned halves (92%) when the face race in the composites was incongruent, $F(1, 45) = 48.87$, $p < .001$, $\eta^2_p = .52$. The reverse pattern was found when the face race was congruent, where accuracy for the aligned halves was higher (96%) than for misaligned halves (94%), $F(1, 45) = 13.11$, $p < .001$, $\eta^2_p = .23$.

When participants classified the race of the Top half, the Race Congruence × Alignment effect was only marginally significant, $F(1, 45) = 3.10$, $p = .085$, $\eta^2_p = .06$. The trend was due to higher accuracy for misaligned halves (98%) than for aligned halves (97%) when face race was incongruent, $F(1, 45) = 4.61$, $p = .037$, $\eta^2_p = .09$, but no alignment effect was found when face race was congruent, $F(1, 45) = 0.02$, $p = .902$, $\eta^2_p < .01$. Hence the three-way interaction between Race Congruence, Alignment, and Position was due to a more pronounced Race Congruence × Alignment effect for classifying the Bottom than for classifying the Top half of the composite stimuli.

Discussion

The main finding of Experiment 2 was that when top and bottom parts of aligned stimuli were of different face races, participants were slower to make a race categorization judgment than when the same parts were presented in a misaligned format. This is consistent with the overall pattern of results in Experiment 1, where the task-relevant gender congruence between the top and bottom parts of a composite face modulated the classical composite face effect. Now, in Experiment 2 we see that the task-relevant race congruence influences the composite effect, with slower responses to race incongruent stimuli in the aligned condition.

Another finding consistent with Experiment 1 was the larger effect of alignment when the task involved judging the bottom part of the image. However, unlike Experiment 1, in

DOI: 10.1177/1747021819899531

Experiment 2 we only found an effect of congruence from the task-relevant category when the bottom part was being categorized. Relative to the bottom part, the top part was judged more quickly, with response times more than 100 ms faster than for categorizing the bottom halves. The lack of a composite effect in the top half condition may therefore simply reflect the relative ease of judging the race from the top halves of the faces; it seems that particularly salient cues to ethnicity are present in the top part of the face alone.

Also similar to the pattern noted from Experiment 1, the task-irrelevant social category of gender congruence now had little influence on the composite effect. This is a striking finding because Gender Congruence had such a strong influence when it was task relevant in Experiment 1. However, we should note the borderline Gender Congruence × Alignment interaction despite its small effect size, $p = .092$, $\eta^2_p = .06$. A similar marginal effect of gender congruence was also reported by Chen et al. (2018) when the task was recognizing familiar face identity. This may suggest that gender is a particularly salient social category that is difficult to ignore completely.

As was noted for the data from Experiment 1, cultural differences (the Participant Race factor) did not interact with the composite effect (the Alignment variable) in Experiment 2. Although we did note a complex four-way interaction involving Participant Race, Race Congruence, Gender Congruence, and Position, this did not influence the composite face effect as such.

General Discussion

Across two separate experiments we contrasted the effects of task-relevant and task-irrelevant visual properties on the composite face effect. Both experiments used exactly the same stimuli, involving orthogonal combinations of top and bottom face parts with congruent or incongruent male and female gender and with congruent or incongruent Asian and Caucasian racial appearance. Only the task was changed to manipulate task-relevance and task-irrelevance. In Experiment 1 a gender categorization task made gender congruence task-relevant and race congruence task-irrelevant, whilst in Experiment 2 a race categorization task made gender congruence now task-irrelevant and race congruence task-relevant.

DOI: 10.1177/1747021819899531

The main findings were straightforward. Using the difference between response times to the same face parts when presented in aligned (i.e. face-like) or in misaligned (not face-like) formats as an index of the composite effect, there was always a strong influence of the task-relevant manipulation. That is, responses to aligned stimuli with incongruent gender in the top and bottom halves were slower than responses to equivalent misaligned stimuli when the task was gender categorization, and responses to aligned stimuli with incongruent race in the top and bottom halves were slower than responses to equivalent misaligned stimuli when the task was race categorization. Although these effects were mainly found for the relatively difficult to judge bottom halves, the overall influence is clear.

In contrast, the task-irrelevant variable did not exert a strong influence on the composite effect in either experiment. We did note a trend toward an influence of gender congruence in the race categorization task, but it was small at best. Much more striking was the way in which the influence of gender congruence or race congruence seemed to be largely determined by task relevance despite both gender and race being salient visual properties of faces. We consider this unlikely to reflect differential interference at the response stage because the critical contrast always involves a comparison between aligned and misaligned stimuli that are themselves created from exactly the same face parts. Any effects due to automatic responses to the face parts themselves are therefore cancelled out by the experimental design. The over-riding difference between the aligned and misaligned stimuli is whether they create a face-like appearance (aligned format) or just look like two separate face parts (misaligned format).

Our findings parallel Calder et al.'s (2000) findings of independence between composite effects for face identity and expression. Taken together with Chen et al.'s (2018) demonstration of a much more limited influence of the identity-specific semantic property of occupation on the composite effect, they establish a strong case for the influence of a relatively early perceptual locus over and above any contributions of other factors since it seems that it is the integration of visual properties that determine gender, race, identity and expression in the aligned stimuli that is driving the composite effect, rather than relatively conceptual properties such as occupation.

One way to think about this is that the aligned condition leads to a perceptual

DOI: 10.1177/1747021819899531

representation that will have an overall male appearance when created from male face parts, an overall female appearance when created from female face parts, and a more gender-ambiguous appearance when created from an incongruent combination of male and female parts. The data from Experiment 1 show that determining the gender of the constituent parts is influenced by the perceived gender of this overall combination. Likewise, the perceptual representation in the aligned condition will have an overall Asian appearance when created from Asian face parts, an overall Caucasian appearance when created from Caucasian face parts, and a more racially-ambiguous appearance when created from an incongruent combination of Asian and Caucasian parts. Experiment 2 shows that determining the race of the constituent parts is influenced by the perceived race of this overall combination.

That said, the importance of attentional factors does still need to be taken into account. Although gender and race are both salient visual properties of faces (Bruce & Young, 2012; Kramer et al., 2017) it is striking how well participants can attend to whichever characteristic is task-relevant whilst largely ignoring the impact of incongruency in a task-irrelevant characteristic. This is consistent with Cheng et al.'s (2018) recent findings indicating a mix of serial and parallel processing. An analysis of the composite effect as *solely* reflecting the creation of an integral holistic representation itself does not seem readily able to deal with this flexibility. If the creation of an integral holistic representation was the sole determinant of the composite effect, it would be expected that incongruent task-irrelevant characteristics would inevitably interfere with decisions about a task-relevant characteristic, because these task-irrelevant properties would by definition be bound together with task-relevant properties in the holistic representation itself. The lack of strong evidence of an influence of task-irrelevant characteristics in either Experiment 1 or Experiment 2 thus implies that additional factors are in play.

This does not of course mean that the creation of a holistic representation is irrelevant; there are multiple lines of evidence that point to the importance of holistic representations for the composite effect (see Rossion, 2013). However, our findings show that the composite effect also involves targeted processing of an attended visual characteristic. In our experiments the visual properties of either gender or race were either selected or deselected

DOI: 10.1177/1747021819899531

for their task-relevance by a top-down attentional modulation that was mainly a controlled rather than an automatic process. Ability to attend selectively to gender or race was found despite other evidence pointing to their relatively automatic interaction across multiple levels in the hierarchy of perceptual processing (Freeman & Ambady, 2011; Ito & Urland, 2003; Johnson, Freeman & Pauker, 2012).

We note that participant race did not significantly modify the nature of the composite effect in our experiments. Despite the substantial cultural differences between China and Britain, and despite widespread claims that Chinese and British participants perceive faces differently and even attend preferentially to different face regions (Blais et al., 2008; Jack et al., 2009, 2012; Wheeler & Kim, 1997), we found no interactions that involved both the Alignment and Participant Race variables. Of course, for reasons already stated we were unable simultaneously to look at other-race effects, and these may be worth exploring in future as they are often thought to relate to holistic processing of faces (Hayward et al., 2013; Michel et al., 2006; Rossion, 2011). At present, though, our results are consistent with other studies that find only modest overall differences between Chinese and British participants in perceptual tasks (Sutherland et al., 2018; Yan, Andrews & Young, 2016; Yan, Andrews, Jenkins & Young, 2016; Yan, Young & Andrews, 2017).

Much of the recent interest in the composite face effect derives from the possibility of using it as a measure of individual differences in holistic face perception. As we noted, however, the paradigms used are at present based around measures that involve interference between holistic perception and attentional and response-based effects. This has led to considerable debate about how best to separate their relative contributions (e.g. Richler & Gauthier, 2013, 2014; Richler, Gauthier, Wenger, & Palmeri, 2008; Rossion, 2013). The importance of such factors is evident from studies linking the composite effect to phenomena involving selective attention (Chua & Gauthier 2015; Chua, Richler & Gauthier 2015; Fitousi 2015, 2016) or object-based attention and perceptual grouping (Curby, Entenman & Fleming, 2016; Curby, Goldstein & Ritter, 2013; Retter & Rossion, 2015). Our findings are consistent with this new apprecation of the difficulty of achieving a workable measure, but they offer the possibility that tasks based on highly salient perceptual categories may provide a promising way forward.

DOI: 10.1177/1747021819899531

DOI: 10.1177/1747021819899531

References

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia, 49*, 2541-2552.

Baudouin, J-Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception, 35*, 531-540.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One, 3*, e3022.

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

Bruce, V., & Young, A. (2012). *Face perception.* Hove, East Sussex: Psychology Press.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 645-651.

Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research, 41*, 1179-1208.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configurational information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 527-551.

Chen, F. F., Jing, Y., & Lee, J. M. (2012). "I" value competence but "we" value social competence: the moderating role of voters' individualistic and collectivistic orientation in political elections. *Journal of Experimental Social Psychology, 48*, 1350–1355.

Chen, W., Ren, N., Young, A. W., & Liu, C. H. (2018). Interaction between social categories in the composite face paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition, 44*, 34-49.

Cheng, X. J., McCarthy, C. J., Wang, T. S. L., Palmeri, T. J., & Little, D. R. (2018). Composite faces are not (necessarily) processed coactively: a test using systems factorial technology and logical-rule models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Chua, K. W., & Gauthier, I. (2016). Category-specific learned attentional bias to object parts. *Attention, Perception, & Psychophysics, 78*, 44-51.

Chua, K. W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention

DOI: 10.1177/1747021819899531

to parts. *Journal of Experimental Psychology: General, 144*, 723-729.

Curby, K. M., Entenman, R., & Fleming, J. T. (2016). Holistic face perception is modulated by experience-dependent perceptual grouping. *Attention, Perception & Psychophysics, 78,* 1392-1404.

Curby, K.M., Goldstein, R. R., & Blacker, K. (2013). Disrupting perceptual grouping of face parts impairs holistic face processing. *Attention, Perception, & Psychophysics, 75*, 83–91.

Fitousi, D. (2015). Composite faces are not processed holistically: evidence from the Garner and redundant target paradigms. *Attention, Perception & Psychophysics, 77*, 2037-2060.

Fitousi, D. (2016). Comparing the role of selective and divided attention in the composite face effect: insights from Attention Operating Characteristic (AOC) plots and cross-contingency correlations. *Cognition, 148*, 34-46.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*, 247-279.

Hayward, W. G., Crookes, K., & Rhodes, G. (2013). The other-race effect: holistic coding differences and beyond. *Visual Cognition, 21*, 1224-1247.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*, 65-74.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization, 10*, 15–41.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616-626.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: how covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102*, 116-131.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*, 1543–1548.

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences, USA, 109*, 7241–7244.

Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review, 124*, 115-129.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*, 255-260.

McKone, E., & Robbins, R. (2011). Are faces special? In A. J. Calder, G. Rhodes, J. V. Haxby, & M. H. Johnson (Eds.), *The Oxford handbook of face perception*. Oxford: Oxford University Press, 149-176.

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.

Murphy, J., Gray, K. L. H., & Cook, R. (2016). The composite face illusion. *Psychonomic Bulletin & Review, 24*, 245-261.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. P*sychological Review, 108*, 291–310.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin, 128,* 3–72.

Retter, T., & Rossion, B. (2015). Global shape information increases but color information decreases the composite face effect. *Perception, 44*, 511-528.

Richler, J. J., & Gauthier, I. (2013). When intuition fails to align with data: a reply to Rossion (2013). *Visual Cognition, 21*, 254-276.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin, 140*, 1281-1302.

Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: perceptual and decisional components. *Journal of Experimental Psychology: Learning Memory and Cognition, 34,* 328–342.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology, 3*, 1–6.

DOI: 10.1177/1747021819899531

Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic face perception. *Visual Cognition, 21*, 139-253.

Rossion, B., & Michel, C. (2011). An experience-based holistic account of the other-race face effect. In A. J. Calder, G. Rhodes, J. V. Haxby, & M. H. Johnson (Eds.), *The Oxford handbook of face perception*. Oxford: Oxford University Press, 215-243.

Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial first impressions across culture: data-driven modelling of Chinese and British perceivers' unconstrained facial impressions. *Personality and Social Psychology Bulletin, 4*4, 521-537.

Tanaka, J. W., & Gordon, I. (2011). Features, configuration, and holistic face processing. In A. J. Calder, G. Rhodes, J. V. Haxby, & M. H. Johnson (Eds.), *The Oxford handbook of face perception*. Oxford: Oxford University Press, 177-194.

Wheeler, L., & Kim, Y. (1997). What is beautiful is culturally good: The physical attractiveness stereotype has different content in collectivistic cultures. *Personality and Social Psychology Bulletin, 23*, 795–800.

Yan, X., Andrews, T. J., & Young, A. W. (2016). Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 423-440.

Yan, X., Andrews, T. J., Jenkins, R., & Young, A. W. (2016). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. *Quarterly Journal of Experimental Psychology, 69*, 1247-1254.

Yan, X., Young, A. W., & Andrews, T. J. (2017). Differences in holistic processing do not explain cultural differences in the recognition of facial expression. *Quarterly Journal of Experimental Psychology, 70*, 2445-2459.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747-759.

Young, A. W. (2018). Faces, people and the brain: the 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology, 71*, 569-594.

DOI: 10.1177/1747021819899531

Figure Captions

Figure 1. Illustration of stimuli used in Experiments 1 and 2, with orthogonal manipulation of congruent and incongruent race and gender. A: Aligned composites created from the top and bottom parts of two different faces. B: Misaligned composites, created by misaligning the top and bottom face parts, such that the center of the bottom half was shifted to line up vertically with one edge of the top half. The coding lists the top part of each stimulus, followed by the bottom part, using the following convention: M = male, F = female; A = Asian (Chinese), C = Caucasian, 1 and 2 = model 1 and 2 in each category.

Figure 2. Mean correct reaction time (RT) for categorizing the gender of face halves as a function of Gender Congruence and Alignment in Experiment 1. Error bars represent one standard error of the means.

Figure 3. Mean correct reaction time (RT) for categorizing the gender of face halves as a function of Position and Alignment in Experiment 1. Error bars represent one standard error of the means.

Figure 4. Mean correct reaction time (RT) for categorizing the race of face halves as a function of Race Congruence and Alignment in Experiment 2. A: Results of the Top position. B: Results of the Bottom position. Error bars represent one standard error of the means.

DOI: 10.1177/1747021819899531

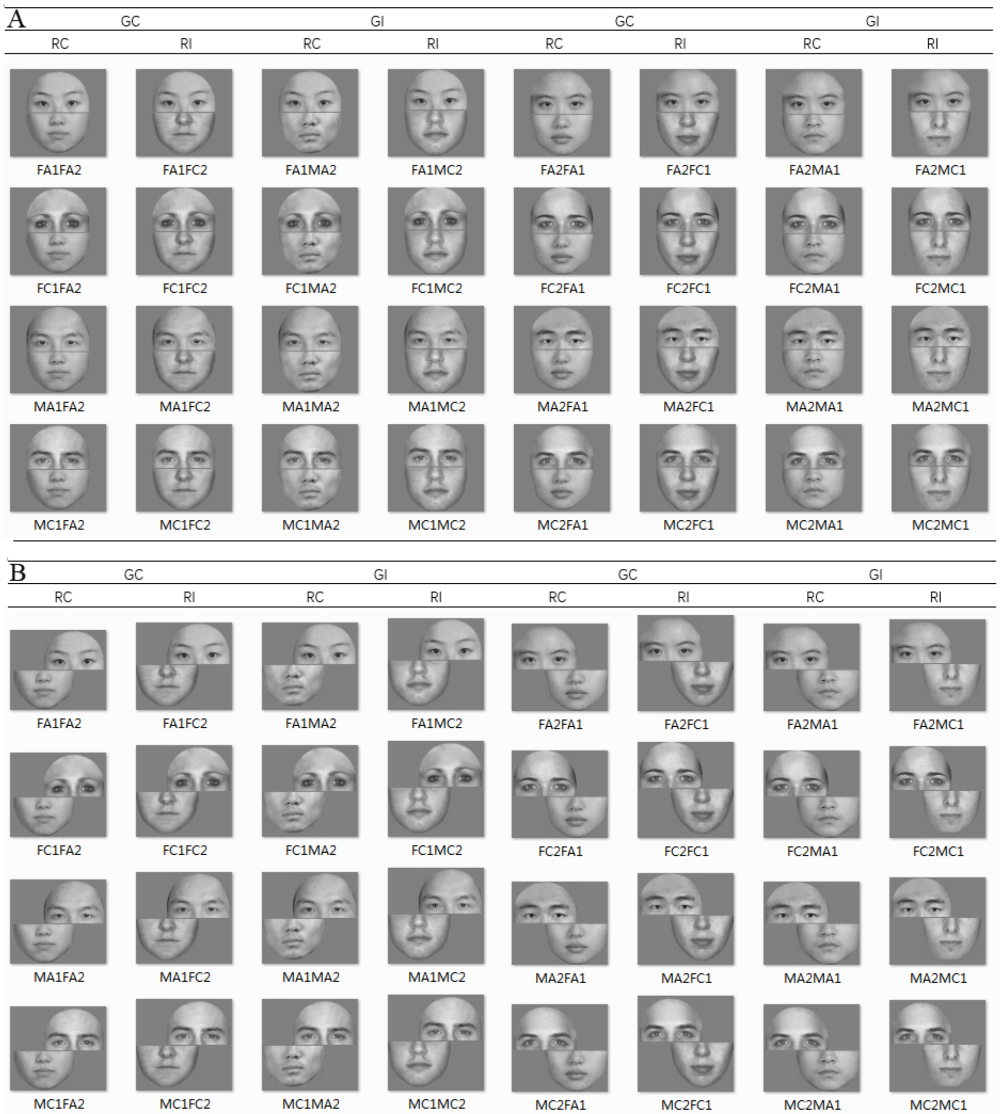

Figure 1. Illustration of stimuli used in Experiments 1 and 2, with orthogonal manipulation of congruent and incongruent race and gender. A: Aligned composites created from the top and bottom parts of two different faces. B: Misaligned composites, created by misaligning the top and bottom face parts, such that the center of the bottom half was shifted to line up vertically with one edge of the top half. The coding lists the top part of each stimulus, followed by the bottom part, using the following convention: M = male, F = female; A = Asian (Chinese), C = Caucasian, 1 and 2 = model 1 and 2 in each category.

DOI: 10.1177/1747021819899531

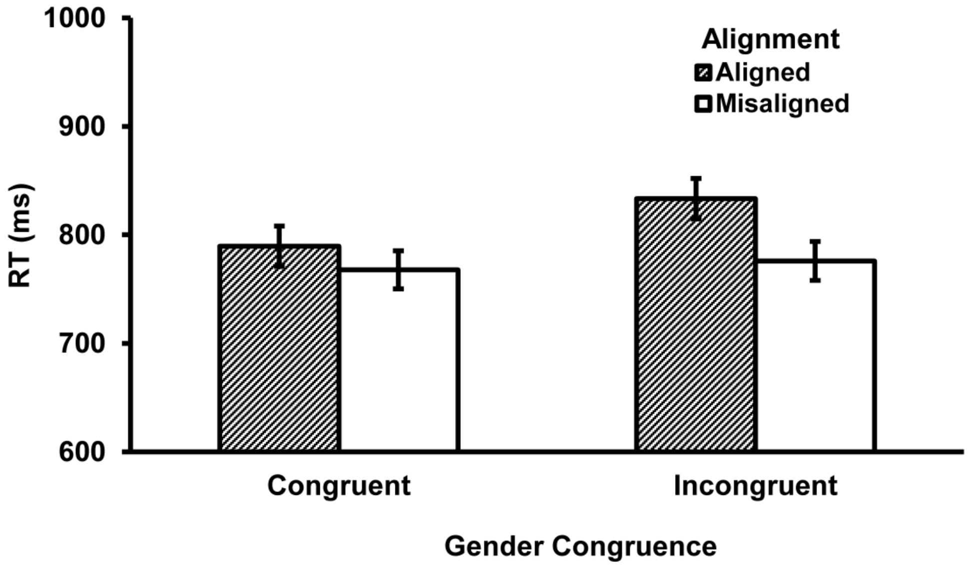

Figure 2. Mean correct reaction time (RT) for categorizing the gender of face halves as a function of Gender Congruence and Alignment in Experiment 1. Error bars represent one standard error of the means.

DOI: 10.1177/1747021819899531

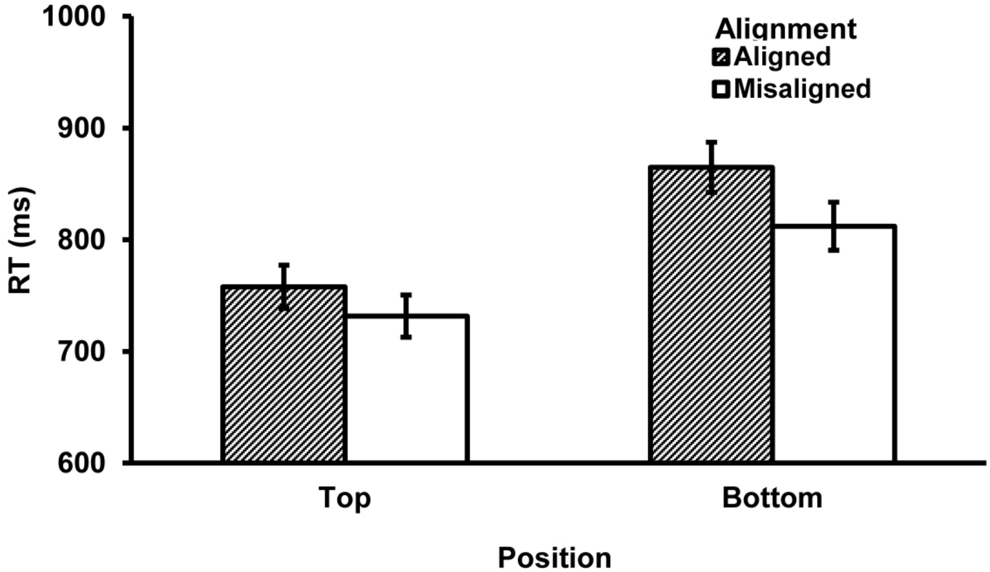

Figure 3. Mean correct reaction time (RT) for categorizing the gender of face halves as a function of Position and Alignment in Experiment 1. Error bars represent one standard error of the means.

DOI: 10.1177/1747021819899531

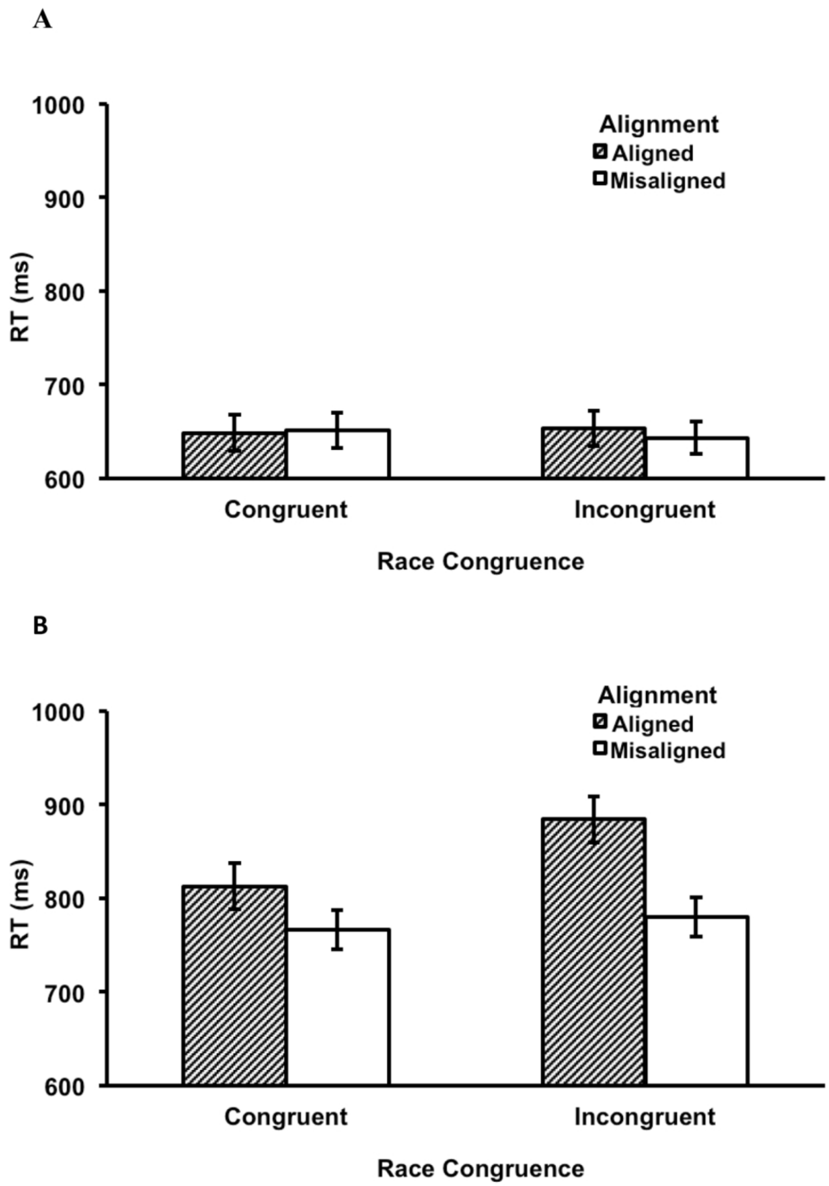

Figure 4. Mean correct reaction time (RT) for categorizing the race of face halves as a function of Race Congruence and Alignment in Experiment 2. A: Results of the Top position. B: Results of the Bottom position. Error bars represent one standard error of the means.

DOI: 10.1177/1747021819899531

Table 1

ANOVA of the Correct Reaction Times for Gender Categorization of Aligned and Misaligned Top and Bottom Face Halves from Experiment 1

Source	*df*	*F*	*p*	η^2_p
Gender Congruence (GC)	1	31.11	<.001	.44
Race Congruence (RC)	1	0.13	.911	<.01
Alignment (A)	1	160.18	<.001	.80
Position (P)	1	19.37	<.001	.33
Participant Race (PR)	1	<0.01	.957	<.01
RC × GC	1	3.64	.064	.09
PR × GC	1	0.95	.336	.02
PR × RC	1	0.29	.591	.01
GC × A	1	15.71	<.001	.29
PR × A	1	0.17	.681	<.01
RC × A	1	1.99	.166	.05
RC × P	1	2.26	.141	.06
GC × P	1	0.02	.883	<.01
PR × P	1	0.08	.774	<.01
A × P	1	7.79	.008	.17
RC × GC × PR	1	0.05	.823	<.01
RC × GC × A	1	0.01	.940	<.01
RC × GC × P	1	0.04	.852	<.01
PR × GC × P	1	0.40	.534	.010
PR × GC × A	1	0.93	.341	.02
PR × RC × A	1	1.38	.248	.03
PR × RC × P	1	0.80	.376	.02
RC × A × P	1	2.03	.162	.05
GC × A × P	1	0.26	.661	.01
PR × A × P	1	1.19	.282	.03
RC × GC × A × P	1	0.18	.671	.01
PR × GC × A × P	1	<0.01	.948	<.01
PR × RC × A × P	1	0.31	.579	.01
GC × RC × PR × P	1	2.13	.153	.05
GC × RC × PR × A	1	2.89	.097	.07
GC × RC × PR × A × P	1	0.35	.555	.01
Error	39			

DOI: 10.1177/1747021819899531

Table 2

ANOVA of the Correct Reaction Times for Race Categorization of Aligned and Misaligned Top and Bottom Face Halves from Experiment 2

Source	*df*	*F*	*p*	η^2_p
Gender Congruence (GC)	1	1.31	.259	.03
Race Congruence (RC)	1	26.63	<.001	.37
Alignment (A)	1	58.38	<.001	.57
Position (P)	1	104.27	<.001	.70
Participant Race (PR)	1	1.34	.253	.02
RC × GC	1	6.21	.016	.12
PR × RC	1	1.53	.223	.03
PR × GC	1	0.18	.676	<.01
RC × A	1	19.09	<.001	.30
PR × A	1	0.38	.541	.01
GC × A	1	2.96	.092	.06
GC × P	1	0.08	.780	<.01
RC × P	1	23.82	<.001	.35
PR × P	1	3.80	.057	.08
A × P	1	101.69	<.001	.69
RC × GC × PR	1	4.96	.031	.10
RC × GC × A	1	0.03	.860	<.01
RC × GC × P	1	0.82	.370	.02
PR × GC × P	1	2.09	.155	.04
PR × GC × A	1	0.87	.357	.02
PR × RC × A	1	<0.01	.947	<.01
PR × RC × P	1	0.05	.833	<.01
RC × A × P	1	9.26	.004	.17
GC × A × P	1	1.04	.314	.02
PR × A × P	1	1.37	.248	.03
RC × GC × A × P	1	1.21	.227	.03
PR × GC × A × P	1	1.01	.322	.02
PR × RC × A × P	1	2.51	.120	.05
GC × RC × PR × P	1	5.76	.021	.11
GC × RC × PR × A	1	1.08	.304	.02
GC × RC × PR × A × P	1	0.02	.901	<.01
Error	45			

DOI: 10.1177/1747021819899531