The composite effect for inverted faces is reliable at large sample sizes and requires the basic face configuration

Tirta Susilo
Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH ✉

Constantin Rezlescu
Department of Psychology, Harvard University, Cambridge, MA ✉

Bradley Duchaine
Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH ✉

The absence of the face composite effect (FCE) for inverted faces is often considered evidence that holistic processing operates only on upright faces. However, such absence might be explained by power issues: Most studies that have failed to find the inverted FCE tested 24 participants or less. Here we find that the inverted FCE exists reliably when we tested at least 60 participants. The inverted FCE was ∼ 18% the size of the upright FCE, and it was unaffected by testing order: It did not matter whether participants did the upright condition first (Experiment 1, $n = 64$) or the inverted condition first (Experiment 2, $n = 68$). The effect also remained when upright and inverted trials were mixed (Experiment 3, $n = 60$). An individual differences analysis found a modest positive correlation between inverted and upright FCE, suggesting partially shared mechanisms. A critical control experiment demonstrates that the inverted FCE cannot be explained by visuospatial attention or other generic accounts because the effect disappeared when the basic face configuration was disrupted (Experiment 4, $n = 50$). Our study shows that the inverted FCE is a reliable effect that requires an intact face configuration, consistent with the notion that some holistic processing also operates on inverted faces.

Introduction

Unlike most objects, upright faces are processed by specialized mechanisms (Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Grand, & Mondloch, 2002; McKone, Kanwisher, & Duchaine, 2007; Rivest, Moscovitch, & Black, 2009; Rossion, 2008; Yin, 1969). Specifically, upright faces are processed by mechanisms that mandatorily integrate information from across an upright face into a single unit of perceptual analysis. This "holistic processing" of upright faces is reflected by many experimental effects including the part-whole effect (Tanaka & Farah, 1993), part-whole in spacing-change effect (Tanaka & Sengco, 1997), categorical perception in noise effect (McKone, Martini, & Nakayama, 2001), superimposed face effect (Martini, McKone, & Nakayama, 2006), and gaze contingent effect (Van Belle, De Graef, Verfaillie, Busigny, & Rossion, 2010).

One effect in particular, namely the face composite effect (FCE, Hole, 1994; Young, Hellawell, & Hay, 1987), has been argued to be the best demonstration of holistic processing for upright faces (Maurer, Grand, & Mondloch, 2002; McKone, Kanwisher, & Duchaine, 2007; Rossion, 2013). The phenomenological basis of the FCE is the face composite illusion (Figure 1). In this illusion, two identical top-halves appear different when they are aligned with different bottom-halves but not when the two halves are misaligned. The illusion reflects holistic processing because top- and bottom-halves that are aligned are perceptually integrated as a single whole face, whereas top- and bottom-halves that are misaligned are not. Building from the illusion, the FCE is traditionally computed as the difference in performance (measured by accuracy and/or response time) when matching identical top-halves in the aligned trials relative to when matching identical top-halves in the misaligned trials. The prediction is that performance should be *worse* for aligned than for misaligned trials because two identical top-halves are likely to be (mistakenly) perceived as different only when they are aligned with different bottom-halves (for a compre-

Figure 1. The face composite illusion. Identical top-halves that are aligned with different bottom-halves appear different (left pair); the illusion disappears when the top- and bottom-halves are misaligned (right pair).

hensive discussion of the rationale behind this comparison see Rossion, 2013). Recently it has been claimed that the FCE should be measured using more experimental conditions (Gauthier & Bukach, 2007; Gauthier & Richler, 2013, but also see McKone & Robbins, 2007; Rossion, 2013), but we limit the scope of the present study to the traditional measure.

The hypothesis that holistic processing operates only for upright faces generates a straightforward prediction, namely that FCE should not be present for objects and inverted faces. Consistent with this prediction, the FCE is absent for objects such as cars (Macchi-Cassia, Picozzi, Kuefner, Bricolo, & Turati, 2009) and dogs (Robbins & McKone, 2007). However, studies of the FCE for inverted faces have found mixed results, although the small samples of these studies (24 participants at most, with the exception of experiments 3 and 4 in McKone et al., 2013) render the issue unclear (Table 1).

The present study was conducted to clarify the existence of the inverted FCE by testing larger samples than did previous studies. A systematic investigation of the inverted FCE will help shed light on the nature of holistic face processing and the extent to which it is exclusive for upright faces. Specifically, a long-standing

controversy in the literature concerns whether upright and inverted faces are analyzed by distinct mechanisms that differ qualitatively (Moscovitch, Winocur, & Behrmann, 1997; Rossion, 2008; Yin, 1969) or by similar mechanisms that differ only quantitatively (Jiang et al., 2006; Sekuler, Gaspar, Gold, & Bennett, 2004; Valentine, 1988).

Here we addressed five questions about the inverted FCE: (1) Does the inverted FCE exist? (2) Does testing order matter for the existence of the inverted FCE? (3) What is the size of the inverted FCE relative to the upright FCE? (4) Do individual differences in inverted FCE correlate with those in upright FCE? (5) Can the inverted FCE be explained by a failure to localize visuospatial attention to the target half (Susilo et al., 2011; McKone et al., 2013) or a generic alignment effect that occurs regardless of stimuli? To address these questions we performed four experiments. In Experiments 1–3 we examined whether the FCE was present for upright and inverted faces in different testing orders: upright then inverted (Experiment 1, $n = 64$), inverted then upright (Experiment 2, $n = 68$), and upright and inverted trials mixed (Experiment 3, $n = 60$). In Experiment 4 ($n = 50$) we tested whether the

Experiment	N	Orientation	FCE (accuracy)	FCE (response time)
Goffaux & Rossion, 2006, experiment 3	21	Between-subjects	Yes	Yes
McKone et al., 2013, experiment 2 (Caucasian Ss)	24	Blocked	No	-
McKone et al., 2013, experiment 2 (Asian Ss)	24	Blocked	No	-
McKone et al., 2013, experiment 3 (caucasian Ss)	32	Blocked	Yes	-
McKone et al., 2013, experiment 3 (Asian Ss)	32	Blocked	Yes	-
McKone et al., 2013, experiment 4 (Caucasian Ss)	92	Blocked	No	No
McKone et al., 2013, experiment 4 (Asian Ss)	22	Blocked	No	No
Mondloch & Maurer, 2008, experiment 1	24	Mixed	No	-
Mondloch & Maurer, 2008, experiment 2	24	Blocked	No	-
Robbins & Coltheart, 2012	24	Between-subjects	Yes	-
Robbins & McKone, 2007, experiment 2	23	Blocked	No	-
Rossion & Boremanse, 2008	18	Mixed	Yes	Yes
Soria-Bauser et al., 2011, experiment 1	20	Blocked	No	No
Susilo et al., 2011, experiment 2	24	Blocked	No	-

Table 1. Previous studies of the FCE for inverted faces. *Notes*: Dashes indicate not examined. The list excludes FCE for familiar faces (e.g., Young et al., 1987) and FCE for face expression (e.g., Calder, Young, Keane, & Dean, 2000).

inverted FCE requires the basic face configuration (i.e., eyes above nose above mouth in a whole face outline).

We tested participants using Amazon Mechanical Turk (www.mturk.com), a popular online testing platform that has been demonstrated to reproduce a variety of classic effects in cognitive psychology (Crump, McDonnell, & Gureckis, 2013). To the best of our knowledge the present study is the first online investigation of the FCE, and this is important for two reasons. First, online studies of face perception and recognition abilities (i.e., how good an individual is in discriminating between faces) have generated reliable data comparable to those collected in the laboratory (Germine et al., 2012). Here we sought to extend these online studies of face *abilities* with online experiments of face *mechanisms* (as captured by the FCE). Second, we wanted to know whether the upright FCE can be reliably obtained in environments less constrained than the laboratory, such as when people are perceiving faces naturally on their computer screen in their homes, offices, coffee shops, etc. Non-laboratory environments are undoubtedly more dynamic and distracting than a well-controlled laboratory setup, but they provide a worthy test of generalization of the FCE and of laboratory studies of face mechanisms more generally.

Experiments 1–3

Method

Participants

We limited participation to individuals who had demonstrated reliable Amazon Turk performance in the past (HIT rate > 95%) and had IP addresses in the United States. A total of 262 individuals took part in our study. Each participant was paid 75 cents for approximately 7–10 minutes of testing. Several participants (7.6%) were excluded from data analysis because they did our experiment more than once or performed very poorly on the task (overall accuracy < 55%).

The final sample consists of 242 participants (127 female) with a mean age of 33.6 years ($SD = 10.9$ years). Experiment 1 tested 64 participants (37 female) with a mean age of 33.5 years ($SD = 11.2$ years). Experiment 2 tested 68 (27 female); mean age was 33.8 years ($SD = 10.5$ years). Experiment 3 tested 60 participants (32 female) who had a mean age of 33.4 years ($SD = 11.1$ years).

Stimuli and procedure

Our stimuli and procedure were adapted from the composite task used in experiment 2 of Susilo et al. (2011). Figure 1 shows example stimuli. Stimuli were made from 60 original faces (32 females), all Caucasian front-view greyscale with neutral expressions. Composite faces were created from original faces with similar skin tone. Lines at the edges of the faces indicated the halves. A black ski-cap was pasted on to cover hair cues. Stimuli subtended 4.04° vertical by 3.05° horizontal (aligned) and 4.04° by 4.61° (misaligned) visual angle when viewed from 80 cm. The final set of stimuli consists of 120 composite pairs (30 same-aligned, 30 same-misaligned, 30 different-aligned, 30 different-misaligned).

On each trial, a pair of composite faces was presented sequentially. The first stimulus appeared for 200 ms, followed by a black screen for 400 ms, and then the second stimulus for 200 ms. Participants had to decide whether the top-halves were the same or different while ignoring the bottom-halves. There were 120 trials per orientation. Six practice trials were provided. Experiment 1 tested the upright condition followed by the inverted condition whereas Experiment 2 tested the reverse. In Experiment 3, upright and inverted trials were randomly mixed.

Data analysis

We computed the FCE in the traditional manner, namely accuracy for same-misaligned trials minus accuracy for same-aligned trials (Le Grand, Mondloch, Maurer, & Brent, 2004; Robbins & Edwards, 2007; Rossion, 2013). We excluded FCEs that fell outside 2.5 standard deviations from the mean FCE in a given condition (at most 7.8% of data). Participants were not instructed to respond quickly, but we analyzed response time data to check for speed-accuracy trade-offs.

Results and discussion

Figure 2 shows the results from Experiments 1–3. In Experiment 1, the upright FCE was significant ($M = 0.22$, $SE = 0.02$, $t_{62} = 12.80$, $p < 0.0001$), replicating the standard finding in the laboratory. The inverted FCE was also significant ($M = 0.03$, $SE = 0.01$, $t_{58} = 3.39$, $p = 0.001$) but was smaller than the upright FCE, $F(1, 63) = 60.41$, $p < 0.0001$. Since it is possible that participants might have been induced to process inverted faces holistically because they did the upright condition first, we tested a new group of participants in Experiment 2 who did the inverted condition first. Both the upright ($M = 0.19$, $SE = 0.02$, $t_{63} = 8.36$, $p < 0.0001$) and the inverted ($M = 0.04$, $SE = 0.01$, $t_{64} = 3.15$, $p = 0.002$) FCEs were significant, with the inverted FCE being smaller than the upright FCE, $F(1, 67) = 54.38$, $p < 0.0001$.
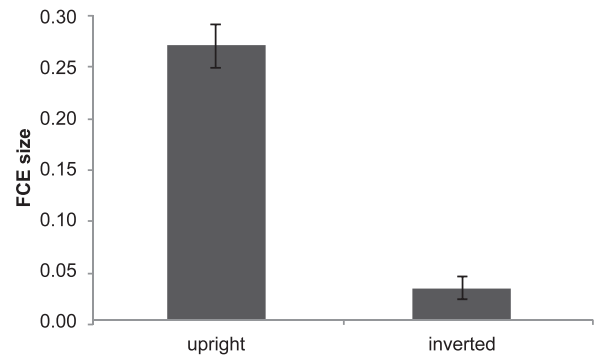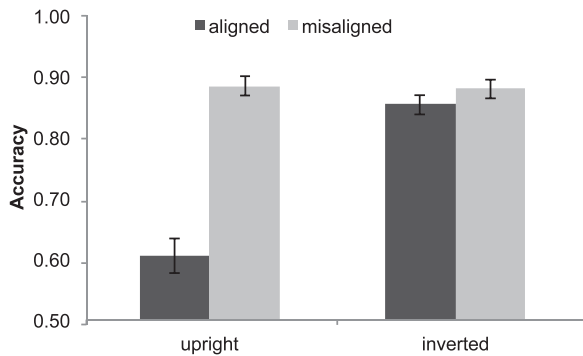
A. Experiment 1
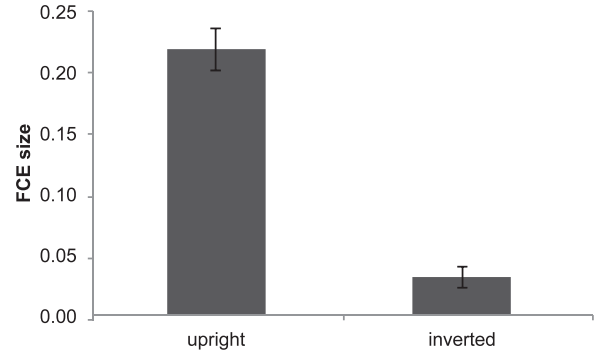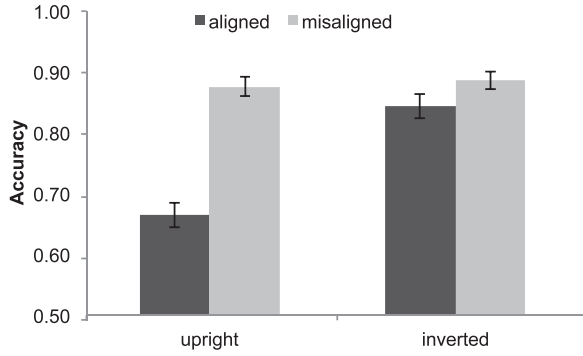
B. Experiment 2
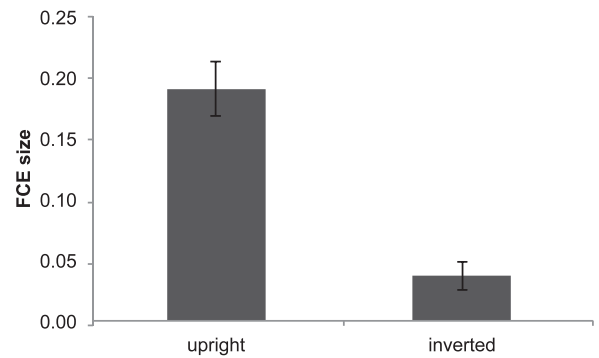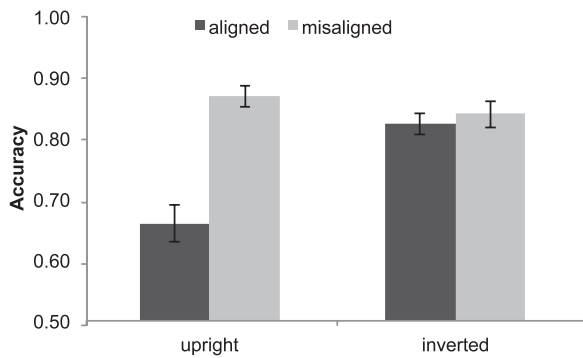
C. Experiment 3

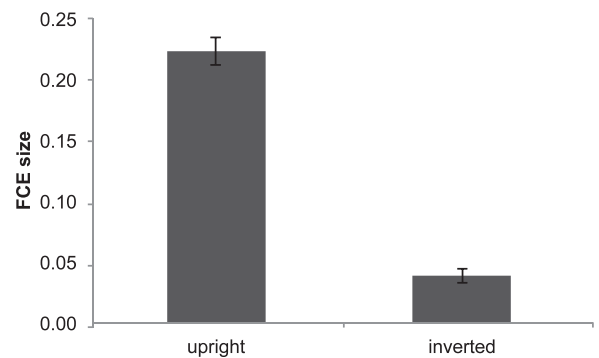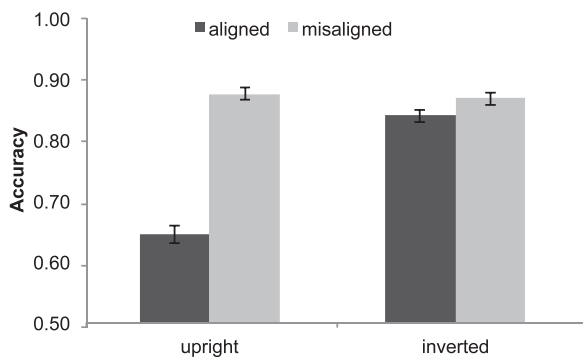

D. Experiments 1-3 (collapsed)

Figure 2. Results of Experiments 1–3. Left panels show accuracy on aligned (dark grey) and misaligned (light grey) trials, whereas right panels show the size of the FCE, for both upright and inverted conditions in Experiment 1 (A), Experiment 2 (B), Experiment 3 (C), and Experiments 1–3 collapsed (D). Error bars show ±1 *SEM*.
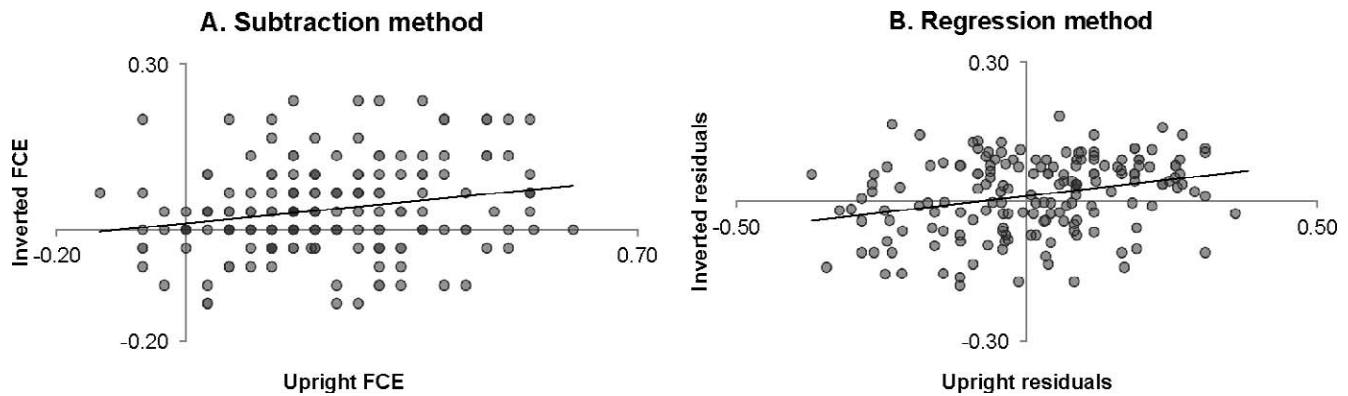
Figure 3. Individual differences analysis of Experiments 1–3 (collapsed). Upright and inverted FCE correlated moderately whether computed using the (A) subtraction approach or (B) the regression approach. Darker dots depict overlaps.

The results of Experiments 1 and 2 show that the inverted FCE is a reliable phenomenon that occurs regardless of testing order. However, since our study was conducted online, it is possible that participants might have tilted their head in inverted trials. To rule out this possibility, and to further replicate the effect, we tested another group of participants in Experiment 3 where upright and inverted trials were randomly mixed. Our presentation time of 200 ms made it impossible for participants to tilt their head in accordance to stimulus orientation. Again we found significant upright ($M = 0.27$, $SE = 0.02$, $t_{58} = 12.69$, $p < 0.0001$) and inverted ($M = 0.03$, $SE = 0.01$, $t_{58} = 2.87$, $p = 0.006$) FCEs, and the inverted FCE was smaller than the upright FCE, $F(1, 59) = 121.00$, $p < 0.0001$. Critically, the size of the inverted FCE in Experiment 3 was not statistically different from those obtained in Experiments 1 ($t_{116} = 0.13$, $p = 0.90$) and 2 ($t_{122} = -0.31$, $p = 0.76$). Together, Experiments 1–3 demonstrate that the inverted FCE is a reliable phenomenon that cannot be accounted by testing sequence, trial order, or participant strategy.

Because the results of Experiments 1–3 are similar, we collapsed the data for a more precise estimate of the size of the inverted FCE and to perform a more powerful individual differences analysis. When the data from Experiments 1–3 were collapsed, the upright FCE was significant ($M = 0.22$, $SE = 0.01$, $t_{182} = 19.42$, $p < 0.0001$), and so was the inverted FCE ($M = 0.04$, $SE = 0.01$, $t_{180} = 6.37$, $p < 0.0001$). Obviously, the inverted FCE was smaller than the upright FCE, $F(1, 191) = 214.39$, $p < 0.0001$. Across three experiments, the size of the inverted FCE was about ∼ 18% of the upright FCE. We also analyzed response time for completion; we observed a significant FCE both upright ($M = 108.50$, $SE = 8.91$, $t_{178} = 11.49$, $p < 0.0001$) and inverted ($M = 17.05$, $SE = 4.46$, $t_{160} = 3.49$, $p < 0.001$). The response time analysis rules out the possibility of speed-accuracy trade-offs.

Next we performed an individual differences analysis to investigate whether the inverted FCE is generated by processes distinct from those underlying the upright FCE, or whether inverted and upright FCE share partially common mechanisms. To do so we examined whether individual differences in inverted FCE were correlated with individual differences in upright FCE. If inverted and upright FCE are generated by distinct mechanisms, then no correlation should be observed. But if inverted FCE and upright FCE are driven by partially shared mechanisms, then a positive correlation should emerge.

Figure 3 plots inverted against upright FCE. Measured using the subtraction method (i.e., accuracy for same-misaligned trials minus accuracy for same-aligned trials), inverted and upright FCE correlated at 0.22, $CI_{95}[0.08, 0.35]$. Given the maximum possible correlation of 0.57 (based on the reliability of the two tasks), the corrected correlation was 0.38. Measured using the regression approach (i.e., residuals for same-aligned trials after regressing out same-misaligned trials, DeGutis, Wilmer, Mercado, & Cohan, 2013), inverted and upright FCE correlated at 0.29, $CI_{95}[0.14, 0.40]$. Given the maximum possible correlation of 0.41, the corrected version of this correlation was 0.68. Both methods of analysis converge to indicate a modest positive relationship between inverted and upright FCE, suggesting that some of the mechanisms underlying the upright FCE also contribute to the inverted FCE.

What mechanisms are tapped by the inverted FCE? One possibility is that the inverted FCE does not reflect holistic processing as it is typically conceptualized in the literature (Rossion, 2013). For example, the inverted FCE may reflect a failure to limit visuospatial attention to the target half (McKone et al., 2013; Susilo et al., 2011), or it may be a generic alignment effect that occurs regardless of stimuli. Alternatively, the inverted FCE may reflect holistic processing that is specialized for upright faces but nevertheless operates on inverted

Figure 4. Vertically swapped composite faces, aligned (left) and misaligned (right). The top row shows stimuli from the upright condition; the bottom row shows stimuli from the inverted condition. The forehead-halves and mouth-halves were vertically swapped to disrupt the basic face configuration (i.e., eyes above nose above mouth).

faces (Freiwald et al., 2009). To tease these alternatives apart, we performed Experiment 4, in which we disrupted the basic face configuration (i.e., eyes above nose above mouth) by vertically swapping the position of the top- and bottom-halves (Figure 4). Composite effects (in both orientations) should occur if the visuospatial account is correct or if the composite task produces a generic alignment effect regardless of stimuli. But if the inverted FCE reflects holistic processing, then its presence would require the basic face configuration, and thus the effect should be abolished.

Experiment 4

Participants

Sixty-five individuals took part in Experiment 4. Data from 15 participants were excluded due to poor performance (overall accuracy < 55%), leaving a final sample of 50 participants (31 female). Mean age was 31.7 years ($SD = 10.1$ years). Participants were randomly allocated to three testing conditions: upright then inverted (22 participants), inverted then upright (11 participants), and a mixed condition (17 participants). Participants were instructed to judge whether the forehead-halves were same or different.

Results and discussion

Figure 5 shows the results. No composite effects emerged either upright ($M = -0.01$, $SE = 0.02$, $t_{48} = -0.68$, $p = 0.497$) or inverted ($M = -0.01$, $SE = 0.01$, $t_{47} = -0.61$, $p = 0.543$). There was a main effect of orientation in that the upright condition was harder than the inverted condition, $F(1, 49) = 13.32$, $p < 0.0001$, but no other effects were significant. We do not know why the upright condition was harder than the inverted condition. Perhaps participants are equally good in matching top-halves in both orientations, but
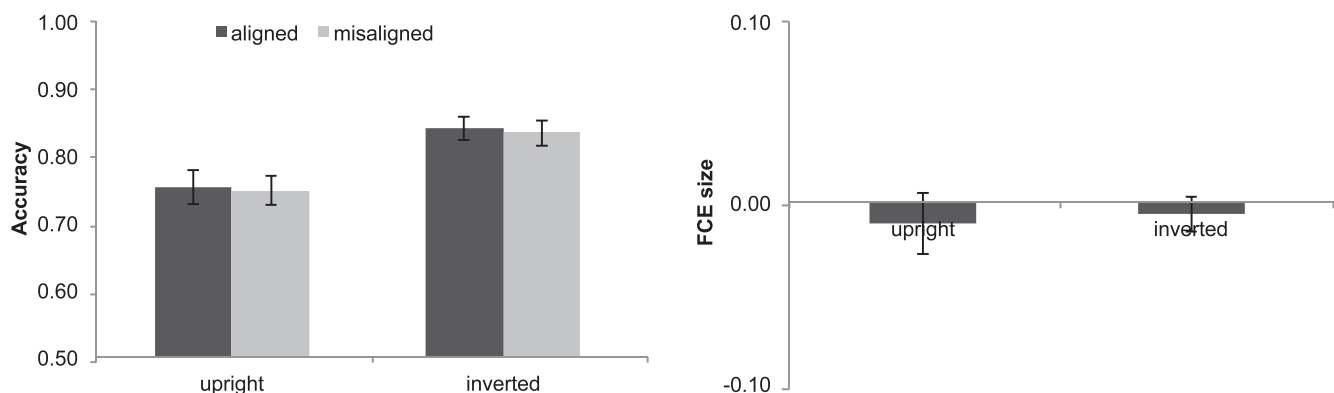

Figure 5. Results of Experiment 4. No difference between aligned and misaligned conditions for both upright and inverted trials (left panel), thus no FCE in either orientation (right panel). Error bars show ± 1 SEM.

in the upright condition they tend to reflexively orient to the upper part of the stimuli (which corresponds to the bottom-half of the face) before attending to the bottom part of the stimuli (which corresponds to the top-half of the face). This potentially interesting finding merits its own investigation.

In summary, Experiment 4 shows that the inverted FCE cannot be accounted by either the visuospatial account or the generic alignment account. There was no composite effect in either orientation when the basic face configuration (i.e., eyes above nose above mouth) was disrupted. This result is consistent with the notion that the inverted FCE is generated by holistic processing mechanisms.

General discussion

The goal of the present study was to address five issues concerning the face composite effect (FCE) for inverted faces: (1) Does the inverted FCE exist? (2) Does testing order matter for the existence of the inverted FCE? (3) What is the size of the inverted FCE? (4) Does the inverted FCE correlate with the upright FCE? (5) Can the inverted FCE be accounted by a failure to localize visuospatial attention to the target half (McKone et al., 2013; Susilo et al., 2011) or by a generic alignment effect that occurs regardless of stimuli? Our findings led us to the following conclusions: (1) The inverted FCE is a reliable phenomenon at large sample sizes; (2) The inverted FCE is robust to testing order; (3) The inverted FCE is ∼ 18% of the upright FCE in size; (4) The inverted FCE correlates modestly and positively with the upright FCE; and (5) The inverted FCE cannot be explained by either the visuospatial or the generic alignment account because its presence requires the basic face configuration (i.e., eyes above nose above mouth in a whole face outline).

Our results contribute to current discussions about the inverted FCE and what it means for theories of holistic face processing. McKone et al. (2013)—and prior to that Susilo et al. (2011)—proposed the visuospatial hypothesis to account for the inverted FCE. According to this hypothesis, the inverted FCE occurs because participants sometimes fail to restrict their visuospatial attention to the target half, leading to a "false" composite effect for inverted faces. McKone et al. further argued that this visuospatial failure is caused by complex interactions between many factors (sex, race, stimulus quality, location of eyes, etc.). Here we identify two factors not discussed by McKone et al. The first factor is sample size. As shown in Table 1, with the exception of one experiment, eight previous experiments that failed to find the inverted FCE tested 24 participants or less. In contrast, our study observed the inverted FCE in three experiments that each tested at least 60 participants. The second factor is the presence of the basic face configuration. In Experiment 4 we directly examined the visuospatial hypothesis by testing whether a "false" composite effect would still occur when the face configuration is disrupted. The absence of composite effect in Experiment 4 is inconsistent with the visuospatial hypothesis. We note, however, that our stimuli in Experiment 4 are not as coherent in their global form as inverted faces and many real world objects are. Thus it is important for future studies to test, with large sample sizes, whether the composite effect also occurs for objects that are naturally perceived as a whole unit (dogs, cars, etc.).

Do our results imply that inverted faces are processed holistically like upright faces, and that processing of upright and inverted faces differ only quantitatively? That the basic face configuration is required to obtain the inverted FCE suggests that inverted faces are processed by at least some of the same mechanisms that generate holistic processing for upright faces, albeit to a much lesser extent (given the small size of the inverted FCE). This finding is consistent with the idea that any stimulus that has been classified as a face by some gating mechanisms will be analyzed by the same perceptual computations (Tsao & Livingstone, 2008). Moreover, face-selective neurons in the macaque middle face patch represent face features in inverted faces in the same way as they do features in upright faces, namely relative to an upright face template (e.g., neurons that represent eyes in an upright face are activated by mouth in an inverted face, Freiwald, Tsao, & Livingstone, 2009). These considerations suggest that holistic face processing may not be best conceptualized as mandatory mechanisms that operate only on upright faces, but rather as perceptual computations that operate on any stimulus that fits an upright face template (including inverted faces).

However, the evidence that holistic processing operates only on upright faces comes not just from the FCE but also from many other experimental effects (McKone & Robbins, 2011). Future studies should explore this issue by testing for the presence of these other effects in inverted faces with large sample sizes. Testing with large sample sizes is important because our study suggests that previous failures to find reliable inverted FCE are likely due to small sample sizes. Post-hoc power analyses show that the probabilities of detecting the inverted FCE in the present study were 90% (Experiment 1), 88% (Experiment 2), and 71% (Experiment 3). Had we tested only 24 participants, the probabilities would have been 52%, 47%, and 35%. In contrast, to detect the upright FCE with 90% power, one would need to test only eight participants. It is also important to note that the FCE is built upon the existence of a robust perceptual illusion (Figure 1). The

illusion is nowhere near compelling for inverted faces (readers are invited to flip Figure 1 upside-down), but then again the inverted FCE is only ~ 18% in size compared to the upright FCE.

Finally, our study demonstrates the viability of conducting face-composite experiments via the web. Similar to a recent study showing the online reproducibility of many classic findings in cognitive psychology (Crump et al., 2013), our study shows that holistic face processing can be studied using online participants. The size of the upright FCE in the present study (19–27%) is in agreement with those obtained in the laboratory using identical stimuli (19% in experiment 2 of Susilo et al., 2011; 16% in experiment 2 of McKone et al., 2013). Our study extends previous online studies that thus far have focused only on face abilities (Germine, Duchaine, & Nakayama, 2011; Susilo, Germine, & Duchaine, 2013; Wilmer et al., 2010) rather than face mechanisms. Our study also attests to the generalizability of the FCE and holistic face processing more generally by showing that the FCE can be obtained outside of the psychological laboratory and that holistic face processing can be investigated in less constrained settings.

Keywords: face, composite, inversion, perception, holistic, configuration

Acknowledgments

References

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

Crump, M. J. C., Mcdonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE, 8*(3), e57410.

DeGutis, J., Wilmer, J. B., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition, 126,* 87–100.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J.

N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498.

Freiwald, W., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience, 12*(9), 1187–1196.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103*(2), 322–330.

Gauthier, I., & Richler, J. (2013). When intuition fails to align with data: A reply to Rossion (2013). *Visual Cognition, 21*(2), 254–276.

Germine, L., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition, 118*(2), 201–210.

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review, 19*(5), 847–857.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial" - holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1023–1039.

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

Jiang, X., Rosen, E., Zeffiro, T., Vanmeter, J., Blanz, V., & Riesenhuber, M. (2006). Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron, 50*(1), 159–172.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*(11), 762–768.

Macchi-Cassia, V., Picozzi, M., Kuefner, D., Bricolo, E., & Turati, C. (2009). Holistic processing for faces and cars in preschool-aged children and adults: Evidence from the composite effect. *Developmental Science, 12*(2), 236–248.

Martini, P., McKone, E., & Nakayama, K. (2006). Orientation tuning of human face processing estimated by contrast matching in transparency displays. *Vision Research, 46*(13), 2102–2109.

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

McKone, E., Davies, A. A., Darke, H., Crookes, K., Wickramariyaratne, T., Zappia, S., … , Fernando, D. (2013). Importance of the inverted control in measuring holistic face processing with the com-

posite effect and part-whole effect. *Frontiers in Psychology, 4*(February), 33.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*(1), 8–15.

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception & Performance, 27*(3), 573–599.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition, 103*(2), 331–336.

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

Mondloch, C., & Maurer, D. (2006). The effect of face orientation on holistic processing. *Perception, 37,* 1175–1186.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9,* 555–604.

Rivest, J., Moscovitch, M., & Black, S. (2009). A comparative case study of face recognition: The contribution of configural and part-based recognition systems, and their interaction. *Neuropsychologia, 47*(13), 2798–2811.

Robbins, R., & Coltheart, M. (2012). Left-right holistic integration of human bodies. *Quarterly Journal of Experimental Psychology, 65,* 1798–1819.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103*(1), 34–79.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*(2), 274–289.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21*(2), 139–253.

Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of Vision, 8*(4):

3, 1–13, http://www.journalofvision.org/content/8/4/3, doi:10.1167/8.4.3. [PubMed] [Article]

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396.

Soria-Bauser, D. A., Suchan, B., & Daum, I. (2011). Differences between perception of human faces and body shapes: Evidence from the composite illusion. *Vision Research, 51,* 195–202.

Susilo, T., Germine, L., & Duchaine, B. (2013). Face recognition ability matures late: Evidence from individual differences in young adults. *Journal of Experimental Psychology: Human Perception & Performance, 39(5),* 1212–1217.

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., . . . , Rhodes, G. (2011). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology, 27*(8), 636–664.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology A, 46*(2), 225–245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*(5), 583–592.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31,* 411–437.

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

Van Belle, G., De Graef, P., Verfaillie, K., Busigny, T., & Rossion, B. (2010). Whole not hole: Expert face recognition requires holistic perception. *Neuropsychologia, 48*(9), 2620–2629.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., . . . , Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences, USA, 107*(11), 5238–5241.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*(6), 747–759.