The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition

Constantin Rezlescu[1,2], Tirta Susilo[3], Jeremy B. Wilmer[4], & Alfonso Caramazza[1]

[1]*Department of Psychology, Harvard University, USA*
[2]*Institute of Cognitive Neuroscience, University College London, UK*
[3]*School of Psychology, Victoria University of Wellington, NZ*
[4]*Department of Psychology, Wellesley College, USA*

Abstract

Face recognition is thought to rely on specific mechanisms underlying a perceptual bias towards processing faces as un-decomposable wholes. This face-specific "holistic processing" is commonly quantified using three measures: the inversion, part-whole, and composite effects. Consequently, many researchers assume that: i) these three effects measure the same cognitive mechanism(s), and ii) these mechanisms contribute to the wide range of individual differences seen in face recognition ability. We test these assumptions in a large sample (N=282), with individual face recognition abilities measured by the well-validated Cambridge Face Perception Test. Our results provide little support for either assumption. The small to non-existent correlations between inversion, part-whole, and composite effects (correlations between -.03 and .28) fail to support the first assumption. As for the second assumption, only the inversion effect moderately predicts face recognition (r = .42); face recognition was weakly correlated with the part-whole effect (r = .25) and not correlated with the composite effect (r = .04). We rule out multiple artifactual explanations for our results by (i) using valid tasks that produce standard effects at the group level, (ii) demonstrating that our tasks exhibit psychometric properties suitable for individual differences studies, and (iii) demonstrating that other predicted correlations (e.g., between face perception measures) are robust. Our results show that inversion, part-whole, and composite effects reflect distinct perceptual mechanisms, and argue against the use of a single, generic term "holistic processing" when referring to these effects. Our results also question the contribution of these mechanisms to individual differences in face recognition.

Statement of public significance

Face recognition is thought to rely on specific cognitive mechanisms that process faces as un-decomposable wholes. This face-specific "holistic" processing is typically quantified using three measures: the inversion, the part-whole, and the composite effect. The implicit assumption is that these effects index the operation of the same unique mechanisms that make significant contributions to face recognition ability. We tested this assumption in a large individual differences study (N =202). First, we show that the inversion, part-whole, and composite effect reflect distinct perceptual mechanisms. Second, we find that face recognition was moderately predicted (r^2 = .18) only by the inversion effect; the part-whole and the composite effect were weak (r^2 = .06) or not significant (r^2 = .00) predictors, respectively. Our study suggests that "holistic" processing is not a single, uniform construct, and that it does not contribute to individual differences in normal face recognition as much as previously considered.

1. Introduction

One of the most influential ideas in face perception research is holistic processing, the notion that there are special cognitive mechanisms that process faces as a perceptual whole or gestalt, without decomposition into smaller visual elements (Tanaka & Farah, 1993). Holistic processing is proposed to explain why humans are good at discriminating between many faces despite their similarity as homogenous visual forms (i.e., all faces are essentially an oval-shaped structure that contains two blobs above an inverted triangle above another blob). As observed by Francis Galton (1883): *"The differences in human features must be reckoned great, inasmuch as they enable us to distinguish a single known face among those of thousands of strangers, though they are mostly too minute for measurement. At the same time, they are exceedingly numerous. The general expression of a face is the sum of a multitude of small details, which are viewed in such rapid succession that we seem to perceive them all at a single glance."* This idea is consistent with several lines of findings. Lack of or diminished holistic processing has been found in individuals with poor face recognition, such as acquired and developmental prosopagnosics (Busigny, Joubert, Felician, Ceccaldi, & Rossion, 2010; Palermo et al., 2011; DeGutis, Cohan, Mercado, Wilmer, & Nakayama, 2012) and individuals with autism spectrum disorders (Hobson, Ouston, & Lee, 1988; Joseph & Tanaka, 2003), suggesting holistic processing is a necessary hallmark of normal face recognition. Reduced holistic processing can also be related to worse recognition of subpopulations of faces with which one has less experience, such as faces from other race, ethnic, or age groups (Susilo, Crookes, McKone, & Turner, 2009; Michel, Rossion, Han, Chung, & Caldara, 2006; but see Hayward, Crookes, & Rhodes, 2013). Moreover, neural responses that are selective for faces as a whole have been identified by electrophysiology (Desimone, Albright, Gross, & Bruce, 1984) and neuroimaging (Kanwisher, McDermott, & Chun, 1997; Liu, Harris, & Kanwisher, 2010; Schlitz & Rossion, 2006). Significantly, the presence of the upright whole face influenced the tuning of neurons in the macaque middle face patch to individual face features (Freiwald, Tsao, & Livingstone, 2009).

Three experimental paradigms have come to be widely regarded as standard measures of holistic processing, as illustrated in Figure 1. One is the *inversion task*, in which turning stimuli upside-down impairs face recognition more than object recognition (Yin, 1969). The disproportionate inversion cost for faces presumably arise because while upright faces are perceived holistically, inverted faces rely on piecemeal processing. Another is the *part-whole task*, in which perception of a face part such as eyes is more accurate when that part is shown within an upright face than when it is presented alone (Tanaka & Farah, 1993). This effect was interpreted to indicate that face parts are encoded not in isolation but rather as integral

elements of a larger visual unit that is an upright face. The third experimental paradigm is the *composite task*, in which perception of the top-half of an upright face is influenced by a perfectly aligned bottom-half even when participants are told to disregard the bottom-half. The perception of the top half is dramatically less affected by the bottom half when the halves are misaligned breaking the wholeness of the face (Young, Hellawell, & Hay, 1987), or when the two halves are aligned but presented upside-down (Susilo, Rezlescu, & Duchaine, 2013). The composite effect was proposed to show that the two face halves are automatically integrated when they are aligned to fit an upright face template.

There are two widely shared (and far-reaching) assumptions permeating research related to the inversion, composite and part-whole effects. The first assumption is that these effects measure the same holistic processes. Several findings showing that the three effects associate across experimental manipulations seem to support this view. The inversion, composite and part-whole effects are greatly reduced or abolished for a wide range of non-face objects and other control stimuli, thus dissociating face recognition from object recognition more generally (McKone & Robbins, 2011). Furthermore, all these effects are found in young children, with adult-like sizes at the youngest age tested – between three and five years old (McKone, Crookes, Jeffery, & Dilks, 2012). That inversion, part-whole, and composite effects are widely assumed to tap the same holistic processes are evident in many review articles on face perception (Duchaine & Yovel, 2008, McKone, Kanwisher, & Duchaine, 2007; Piepers & Robbins, 2012; Rhodes, 2013; Tanaka & Gordon; 2011), yet this assumption is seldom tested (DeGutis, Wilmer, Mercado, & Cohan, 2013, Wang, Fang, Tian, & Liu, 2012).

The second assumption is that the inversion, part-whole and composite effects, as putative measures of holistic processing, are related to face recognition ability, in that larger effects (i.e. stronger holistic processing) are associated with better face recognition. This notion is consistent with findings that inversion, part-whole, and composite effects can be abolished or reduced in individuals with impaired face recognition (Busigny et al., 2010; Palermo et al., 2011; DeGutis et al., 2012; Ramon, Busigny, & Rossion, 2010). However, the relationship between holistic processing and face recognition ability[1] in the normal range remains unclear (DeGutis et al., 2013, Konar, Bennett, & Sekuler, 2010; Richler, Cheung, & Gauthier, 2011a; Wang et al., 2012).

---

[1] In this paper we use face recognition as a general term for the ability to process the identity of faces. Face recognition has at least a perceptual component (that allows us to discriminate between two faces) and a memory component (responsible for matching a current face representation with a previously stored representation). We call these components face perception and face memory.

The goal of the present study was to test the assumptions that inversion, composite, and part-whole effects reflect common underlying mechanisms and that these effects are indicative of face recognition ability. To do so we ask whether and to what extent these effects associate across individual variations in the normal range. Several studies (DeGutis et al., 2013; Konar et al., 2010; Richler et al., 2011a; Wang et al., 2012) have taken a similar approach to examine one or the other assumption, but a series of limitations and inconsistent methods left the status of these questions unclear. First, none of these studies tested inversion, part-whole and composite effects in the same group of participants, preventing a direct comparison of the extent to which each effect is related to each other and to face recognition. Second, with the exception of DeGutis et al. (2013), the researchers used subtraction to compute the effects. An example of subtraction is when the part-whole effect is computed as the difference in performance between parts presented in the context of the whole face ("whole" condition) versus parts presented in isolation ("parts" condition). In individual differences studies, however, subtraction does not properly isolate the effect specific to the condition of interest from the control condition. Instead, subtraction produces measures that correlate with (and are thereby confounded with) the control condition. For example, the part-whole effect, computed via subtraction, may correlate strongly with performance in the parts condition, a condition that is used as a control condition precisely for its theorized lack of holistic processing (DeGutis et al, 2013). Therefore, it is not accurate to refer to a measure that subtracts non-face processing from face processing as "face-specific", as is sometimes done in individual differences studies (e.g. Zhu et al, 2010; Wang et al., 2012). Third, some results were inconsistent. Wang et al. (2012) reported face recognition (when computed as a difference between face and flower perception, but not when reported as an absolute value) was modestly predicted by the standard composite and part-whole effects (r = .13 for both), while the other studies (DeGutis et al., 2013; Konar et al., 2010; Richler et al., 2011a) reported no such correlations. We note that DeGutis et al. (2013) and Richler et al. (2011a) reported significant correlations between face recognition and the size of an alternative effect obtained with composite faces, in a test originally introduced by Farah, Wilson, Drain, & Tanaka (1998) and more recently used by Richler and colleagues (e.g. Cheung, Richler, Palmeri, & Gauthier, 2008; Richler, Gauthier, Wenger, & Palmeri, 2008; Richler et al., 2011a; Richler, Wong, & Gauthier, 2011d). This finding did not replicate in two follow-up studies (Richler et al., 2014; 2015). There has been an extensive, complex debate in the literature about the differences between the alternative effect

and the classic effect based on the standard composite paradigm[2]. Conceptually and by design, the alternative composite test measures a different phenomenon – more akin to attentional interference by incongruent information (Robbins & McKone, 2007; Rossion, 2013). The alternative design adds a number of experimental conditions to explicitly measure the effect on congruency between the correct response (same or different) based on the target face part (e.g. top half) and the status of the task-irrelevant face part (e.g. same or different bottom-half). Importantly, this design produces sizable effects in a wide range of visual stimuli, such as inverted faces (Richler, Mack, Palmeri, & Gauthier, 2011c), novel objects (Gauthier & Tarr, 2002; Gauthier, Williams, Tarr, & Tanaka, 1998; Wong, Palmeri, & Gauthier, 2009), cars (Bukach et al., 2010; Gauthier, Curran, Curby, & Collins, 2003) or words/characters (Wong et al., 2011; Wong et al., 2012), often comparable to the effects obtained for upright faces. Furthermore, a recent review of the alternative composite effect (Richler & Gauthier, 2014) found no correlation with the standard composite effect, additional evidence that the two effects measure distinct phenomena. Therefore, while definitely interesting in the general context of congruency effects (just like the Navon effect, for example; Navon, 1977), the alternative composite design does not capture the kind of upright face-specific mechanisms we were interested in this paper (for readers interested in other examples of general factors correlating with face recognition ability, such as fluid intelligence, short-term memory, lexical knowledge ability, and g, we recommend a recent paper by Gicnac et al., 2016).

Our approach was to replicate and extend these studies by simultaneously testing inversion, composite, and part-whole effects using reliable and valid perceptual tasks that produce standard effects at the group level. Our paper makes four contributions to the literature: (A) it looks at a broad array of putative face holistic processing tests, something that has not been done before; (B) it uses regression rather than subtraction, thereby more cleanly isolating face-specific processing; (C) it uses reliable measures and it considers all of its correlations explicitly in terms of upper bounds set by the reliability of its measures; (D) it measures face recognition abilities using a reliable and well-validated test of face perception with minimum memory demands, the Cambridge Face Perception Test (Duchaine, Germine, & Nakayama, 2007; Bowles et al., 2009). Because holistic processing is conceptualized a perceptual phenomenon, a face perception test is better suited for capturing potential links between holistic processing and face recognition ability than a face memory test.

---

[2] The key arguments of the debate can be found in Robbins & McKone (2007), Richler et al. (2011a, 2011b, 2011c, 2011d), Rossion (2013), Richler & Gauthier (2013, 2014).

2. Method

2.1. Participants.

We tested participants over the web and in the laboratory. Our original web sample consisted of 247 participants from Amazon Mechanical Turk, of which we excluded 45 based on abnormally fast response times suggesting lack of engagement with the task. Our final web sample consisted of 202 participants (80 male) with mean age 36.2 years (SD = 11.5).

Our original lab sample consisted of 84 Harvard psychology students, of which we had to exclude four because of technical problems. Our final lab sample consisted of 80 participants (34 male) with mean age 20.5 years (SD = 2.1).

2.2. Procedure.

Participants completed four experimental tasks that were presented in random order. The first measured face perception abilities, while the other three measured holistic processing.

Cambridge Face Perception Test - CFPT (Duchaine, Germine, & Nakayama, 2007; Duchaine, Yovel, & Nakayama, 2007). In this test participants are asked to sort six test faces (side view) according to how much they resemble a target face (frontal view). Test faces are created by morphing the target face with six identities, and selecting the images with 88%, 76%, 64%, 52%, 40% and 28% of the target face on the morph continua. The test has eight upright trials and eight inverted trials. Participants were allowed 40 seconds per trial (in the original CFPT participants had 60 seconds per trial). Following Rezlescu, Pitcher, & Duchaine (2012), the correct score is calculated as one minus the percentage of total errors. Chance level is 35.6%. The demo version of the experiment can be found at: testable.org/t/404cb3f.

Face Matching Test (Rezlescu et al., 2012). In this test, participants are asked to match one of three test faces (side view) with a target face (frontal view) in terms of identity. All stimuli are grayscale male faces with a standard black cap to cover the hair. Participants first see the target face for 400ms, followed immediately by the three simultaneously presented test images for 2000ms, and have unlimited time to select a response. The chance level is 33.3% and the test has 60 upright trials and 60 inverted trials (randomized presentation). The inversion effect was computed considering the performance on upright and inverted trials using subtraction and regression (see section 2.3. Data analysis). The demo version can be found here: testable.org/t/4b3e988.

Part-Whole Test (DeGutis et al., 2013; Tanaka, Kiefer, & Bukach, 2004). This test presented trials consisting of one target image of a whole face and two test images: either two whole faces ("whole" trials) or two isolated features ("part" trials). The target appeared for

1000ms, followed by a mask (scrambled face) for 500ms, and the test images until participants indicated which of the test stimuli matched the target. Chance level was 50%. The stimuli were created from two Caucasian face templates (one male). For each template, six target faces were created by adding distinct noses, mouths, and eyes. Target faces were unique and did not share any feature between them. Whole foils for each target face were created by switching one of the distinct features (eyes, nose, or mouth) with that of a different target face. The test had 72 upright "whole" trials and 72 upright "part" trials (24 per feature), with an equal number of male and female targets. The same trials were also presented with the stimuli upside-down, for a total of 288 randomly intermixed trials. The part-whole effects were calculated considering the "whole" and "part" trials using subtraction and regression (see section 2.3. Data analysis). The demo version can be found here: testable.org/t/45132d6.

Face Composite Test (adapted from Susilo et al., 2013). Face composites were created by combining the top (i.e. forehead) and bottom (i.e. mouth) halves from 60 faces (32 female) with similar skin tone and neutral expressions. Each composite face had two versions: one aligned, in which the top and bottom halves were perfectly aligned to fit a standard face template, and one misaligned in which the bottom half was shifted to the left or to the right by half the face width. One trial consisted of two composite faces, both aligned or both misaligned, presented sequentially (the first composite for 200ms, a blank screen for 400ms, and the second composite for 200ms). Participants were asked to ignore the mouth halves and say whether the forehead halves of the two composites are the same or different. The mouth halves were always different. The test had 90 upright trials: 30 same-aligned, 30 same-misaligned, 15 different-aligned and 15 different-misaligned. The same trials were also presented with the stimuli upside-down, for an additional 90 inverted trials. All 180 trials were randomized. The composite effects were calculated considering the same-aligned and same-misaligned trials using subtraction and regression (see section 2.3. Data analysis). The different-aligned and different-misaligned trials were ignored (they were filler trials). This is the standard version of the composite test. The demo version can be found here: testable.org/t/44c14e3.

2.3. Main measures of face perception and holistic processing

Our main measure of face perception was the score for the upright condition in the CFPT.

Holistic processing was measured by the inversion, part-whole, and composite effects. Each of these effects has a control (or baseline) condition and a condition of interest, and the effect is computed by comparing the two. In contrast to previous individual differences studies that computed these effects by *subtracting* the score in the control condition from the score in

the condition of interest, we computed these effects by *regressing* the control condition from the condition of interest to more cleanly isolate face-specific processing (DeGutis et al., 2013). In this case, the individual effects are the residuals from the regressions. Specifically, here is how we applied regression to our measures of holistic processing: i) for the inversion effect, we regressed out the score in the inverted condition from the score in the upright condition in the Face Matching Test; ii) for the part-whole effect, we regressed out the score in the part upright condition from the score in the whole upright condition in the Part-Whole Test; iii) for the composite effect, we regressed out the score in the same-misaligned upright condition from the score in the same-aligned upright condition in the Face Composite Test.

2.4. Additional measures of face perception and holistic processing

To ensure the robustness of our findings, we considered additional measures of face perception and holistic processing. The analysis associated with these additional measures is presented as Supplemental Material and reveals similar results to the ones reported in the main text.

For face perception abilities, an additional measure was the upright score in the Face Matching Test. For the inversion effect, an additional measure was computed considering the upright and inverted scores in the CFPT.

For the composite and part-whole effects, we also computed the upright-specific effects by regressing out the effects obtained for the inverted trials from the effects obtained for the upright trials (our main measures).

For all measures, for completion we also computed the effects using subtraction instead of regression (with the caveat mentioned earlier, that subtraction-based measures correlate with the control condition, and therefore do not represent "specific" effects of interest in individual differences studies).

2.5. Statistical significance of differences between correlations

To assess the statistical significance of the differences between correlations we used the web utility developed by Lee & Preacher (2013). The online calculator, with formulas from Steiger (1980), compares two correlation coefficients obtained from the same sample when the two correlations have one variable in common.

3. Results

3.1. Test validation.

We took several steps to ensure our tests are valid for an individual differences investigation.

First, we computed the internal reliability of the four tests as measured using Guttman's $\lambda 2$ (Table 1) and Cronbach's α (Supplemental materials). The reliabilities were good, ranging between .61 and .83 for the upright conditions. Lower reliabilities are to be expected for the inverted conditions because the means are generally closer to the chance level, therefore restricting score variance. Our web and lab samples produced very similar reliabilities for all tests (except the inverted part condition in the part-whole test).

Second, we examined overall performance at the group level (Table 1). The conditions in all four tests had scores in line with previous studies (CFPT and/or inversion effects: Bowles et al., 2009; Duchaine, Germine, & Nakayama, 2007; Duchaine, Yovel, & Nakayama, 2007; Rezlescu et al., 2012; composite and/or part-whole effects: DeGutis et al., 2013; Susilo et al., 2013; Wang et al., 2012), and the means and standard deviations for CFPT and Face Matching ensured measurements of face recognition abilities were not affected by ceiling or floor effects.

Third, we obtained all holistic effects at the group level (composite effect: 27% in the web sample and 24% in the lab sample; part-whole effect: 9% in the web sample and 7% in the lab sample; inversion effects: 19% for CFPT web and 20% for CFPT lab; 27% for Face Matching web and 28% for Face Matching lab; all scores computed by subtraction).

Finally, we confirmed our experimental design was appropriate to reveal potential correlations. In theory, this was ensured by the good reliabilities of the measures to be correlated. A threshold for the maximum possible correlation between two measures can be computed as the geometric mean of their reliabilities, and higher reliabilities leave more room to find existing correlations. We indicate this maximum threshold for all reported correlations (the grey bars in Figures 1 and 2). In practice, the substantial correlation ($r = .52$, $p < .001$) between our measures of face perception abilities – CFPT and Face Matching upright – demonstrates that our study can reveal common mechanisms underlying different test performance.

Based on the very similar reliabilities and overall results in the web and lab samples, we combined the two samples in order to further increase the power to detect meaningful correlations. The results reported from here on assume a single sample of 282 participants. Separate analyses for web and lab participants revealed similar results and are available as supplementary material.

3.2. Correlations between holistic measures.

Figure 1 presents the relationships among the three putative measures of holistic effects: inversion, part-whole and composite. The inversion effect was positively correlated with the part-whole effect (Pearson's r = .28, p < .001, 95% CI = [.17 .38]). Considering that two perfectly correlated measures (i.e., r = 1.0) would generate a correlation of only .49 due to imperfect test reliabilities (Guttman's λ2 for the inversion and part-whole effects were .53 and .45, respectively), the obtained correlation between the inversion and part-whole effects corresponds to an adjusted r of .57 (95% CI = [.35 .78]) in the hypothetical case of perfect test reliabilities.

In contrast, there were no correlations between the inversion and composite effect (r = -.03, p = .669, 95% CI = [-.14 .09]), and between the part-whole and composite effect (r = .05, p = .385, 95% CI = [-.07 .17]). Note that the lack of correlations cannot be accounted by poor test reliabilities; if anything, the chances of finding significant correlations between the composite effect and the other two measures were among the highest (see the grey bars in Figure 2 showing the maximum possible correlations based on test reliabilities). Similarly, the lack of correlations involving the composite effect was not due to a reduced test sensitivity to individual differences in the misaligned condition, in which many participants were close to ceiling (M = .89, SD = .09). When we re-computed the correlations including only participants who scored maximum .90 in the misaligned condition (N = 141; misaligned condition performance: M = .82, SD = .08). The results were the same: the composite effect did not correlate with the inversion (r = .03, p = .736) or the part-whole effect (r = .09, p = .303). Furthermore, we also computed the composite effect as a difference in response times to correct trials in the aligned versus misaligned conditions, as is sometimes done (Rossion, 2013). Again, the results were unchanged: there was no correlation with the inversion (r = .09, p = .143) or the part-whole effect (r = -.04, p = .471) (both correlations based on the whole sample of N = 202).

Critically, the correlation between the inversion and part-whole effects was significantly higher than the correlation between the inversion and composite effects (z-score = 3.86, p < .001) and the correlation between the part-whole and composite effects (z-score = 2.73, p = .006).

3.3. Correlations between face perception and measures of holistic processing.

Figure 2 presents the association of each holistic effect with face perception abilities. Face perception was predicted by the inversion effect (r = .42, p < .001, 95% CI = [.31 .51]; adjusted r = .69, 95% CI = [.52 .85]) and by the part-whole effect (r = .25, p < .001, 95% CI = [.14 .36]; adjusted r = .45, 95% CI = [.25 .65]), although to a lesser degree (z-score = 2.46, p = .014).

In contrast, face recognition was not predicted by the composite effect (r = .04, p = .467, 95% CI = [-.07 .16]; adjusted r = .06, 95% CI = [-.10 .23]). The essentially null correlation between CFPT and the composite effect was significantly different from the previous two correlations (z-score = 4.63, p < .001, when compared to the correlation between CFPT and the inversion effect; z-score = 2.61, p = .009, when compared to the correlation between CFPT and the part-whole effect), and it was not due to poor test reliabilities (see the gray bars in Figure 3). The null correlation between CFPT and the composite effect was not due to potential ceiling effects in the composite misaligned condition. When participants who scored above .90 in the misaligned condition were excluded (new sample size: N = 141), the correlation was r = .04 (p = .606). Similarly, CFPT scores were not predicted by composite effects based on response times (r = .10, p = .11; based on the whole sample of N = 202).

A linear regression with CFPT as the dependent variable and the three measures of holistic processing as independent predictors confirmed the inversion and part-whole effects were significant predictors (inversion effect: standardized coefficient Beta = .38, t = 6.69, p < .001; part-whole effect: Beta = .15, t = 2.60, p = .01) while the composite effect was not (Beta = .05, t = .84, p = .400).

4. Discussion

4.1. Summary of findings.

Our study has two major findings. First, while the inversion and part-whole effects were modestly correlated, the composite effect did not correlate with either. Second, the three measures showed varied relationships with face recognition: face recognition was moderately predicted by the inversion effect, weakly predicted by the part-whole effect, and not predicted by the composite effect.

4.2. The composite, part-whole and inversion effects reflect distinct (holistic?) mechanisms.

One question examined in this study was whether the various putative measures of holistic processing tap the same perceptual mechanisms. The widespread assumption in the

face perception literature is that they do. Many review articles consider the inversion, composite, and part-whole effects different ways to measure the same phenomenon (Behrmann, Richler, Avidan, & Kimchi, in press; Duchaine & Yovel, 2008; McKone et al., 2007; Piepers & Robbins, 2012; Rhodes, 2013; Tanaka & Gordon, 2011). Indeed, studies that use one measure often draw conclusions about holistic processing in general (e.g. Avidan, Tanzer, & Behrmann, 2011; Palermo et al., 2011; Susilo et al., 2010; Pellicano & Rhodes, 2003; Konar et al., 2010; Richler et al., 2011a).

Our results show that there is no common mechanism explaining the three putative effects of holistic processing. The results involving the composite effect are especially clear: zero correlation with the part-whole effect and zero correlation with the inversion effect. The lack of correlation between the composite effect and the part-whole effect is broadly consistent with previous studies (Wang et al., 2012; DeGutis et al., 2013 – note that this study reported a correlation between the part-whole effect and the alternative composite effect, but not with the standard composite effect).

The composite and part-whole effects are the most widely used direct measures of a, presumably, unique construct - holistic processing. Both effects are similar in that they rely on a difference in performance between a condition in which a complete, normal face is present and a condition in which there is no face or the face is anomalous (with the top and bottom halves misaligned). But they also differ in notable ways, which makes them less likely to measure the same phenomenon. First, the part-whole effect is driven by increased *accuracy* in the whole (normal face) condition, while the composite effect relies on increased *errors* in the aligned (normal face) condition. Second, the whole condition in the part-whole effect is essentially a face identification (matching) task, while the composite effect does not have such a component (the whole faces in each trial are always different). Third, in the part-whole task participants may use two strategies to respond correctly, that is by observing a match between the correct test stimulus and the target, or by observing a mismatch between the incorrect stimulus and the target (another strategy is to completely ignore one stimulus). In the composite test however, attention to both stimuli is required and a single same/different judgment is possible. Fourth, the part-whole effect may also reflect a memory component in addition to the perceptual one, because each target is presented twelve times. These differences may explain the null correlation between the composite and part-whole effects and why, although they are proposed to tap the same mechanisms, the composite and part-whole effects are in fact measures of distinct phenomena.

The lack of correlations between the composite and the inversion effects may be hastily interpreted to suggest the mechanisms behind the composite effect are not specific to the upright faces. After all, the inversion effects, by definition (regressing out the inverted from the upright performance), should capture all mechanisms that operate on upright but not inverted faces. Therefore, the inversion effects should correlate with other upright-specific mechanisms. This reasoning is valid under one condition: if we assume that the inversion and composite effects both measure the *efficiency* of face-specific mechanisms. However, we believe the composite effect is not a measure of efficiency, but a hallmark of certain *type* of information entering face computations. This idea, consistent with the much larger composite effect found for upright faces than for inverted faces, is explored more in the following section.

The part-whole effect and the inversion effects were significantly correlated, but the size of the correlations does not necessarily imply they measure the same phenomenon. The correlation (r = .28, adjusted r = .57) corresponds to a shared variance of under 8% (or 32% in the ideal case of perfect test reliabilities). For comparison, in our study the correlation between upright scores in the CFPT and the Face Matching Test, two measures of face perception abilities, was significantly higher (r = .52, adjusted r = .72; adjusted explained variance 52%).

Despite the straightforward evidence against a common set of mechanisms measured by the three putative holistic tests included in our study, our results do not rule out the existence of subtypes of holistic mechanisms. Our findings indicate nevertheless that the term "holistic processing" should be used judiciously when referring to inversion, part-whole and composite effects. Researchers should be explicit that by "holistic processing" they almost always mean whatever is measured by that particular design.

Defining holistic processing strictly in terms of the experimental effects leads to another issue: how can a single construct have (at least) three operational definitions with effects that do not correlate (or modestly correlate)? This highlights a general shortcoming of operational definitions. In fact, some researchers argue against the use of operational definitions on these ground (Green, 1992; Leahey, 1980; Lilienfeld et al., 2015). Hence, in the absence of a clear theoretical definition of holistic processing leading to development of research designs to test it and obtain comparable results, a more sensible suggestion is to drop the use of the term holistic processing altogether and refer strictly to the effects obtained (e.g. composite effect, part-whole effect).

4.3. Individual differences in face perception abilities are not driven by the (holistic?) mechanisms underlying the composite and part-whole effects.

The second question of our study was whether face perception abilities are predicted by any of the three holistic processing measures. Previous studies provided mixed evidence. Konar et al. (2010) reported no significant correlation between the composite effect based on response times and a face identification task. In contrast, Wang et al. (2012) found that a face-matching test showed a small correlation with both the standard composite and the part-whole effect (although the primary analyses in this study, based on subtraction scores, did not isolate face-specific processing). DeGutis et al. (2013) also found a significant correlation between face memory and the part-whole effect, but no correlation with the standard composite effect. Finally, Richler et al. (2011) reported no correlation between face recognition and the standard composite effect[3]. In short, there is insufficient evidence in the literature regarding the relation between direct and face-specific measures of holistic processing and face perception abilities. Furthermore, the relation with the indirect measure of holistic processing (the inversion effect) has not been investigated.

In our study, of all the effects purported to measure holistic processing, the inversion effect had the highest correlation with face perception abilities, evidence that our remarkable ability to recognize faces is dependent on perceptual mechanisms engaged by upright but not inverted faces. This is at odds with a previous claim by Sekuler, Gaspar, Gold, & Bennett (2004) that "upright and inverted face processing differ quantitatively, not qualitatively". Their claim was based on the unsurprising result that when part-based face processing is enforced (by varying noise in different face regions), participants rely on similar strategies to discriminate upright and inverted faces. However, the composite and part-whole effects are strong indicators that upright face perception in normal conditions (i.e. with the whole face visible) goes beyond individual parts processing. Unfortunately, the fact that inversion effects can predict face perception abilities reveals little about the precise nature of the mechanisms performing face recognition.

The two direct measures of holistic processing had limited (in the case of part-whole) or no (for composite) power to predict face recognition. Our results add to the evidence suggesting that, at least for a normal population, increased holistic processing as indexed by these effects does not correspond to better face recognition. This may appear at odds with the posited necessary role of holistic processing in normal face perception (e.g. Busigny et al., 2010;

---

[3] DeGutis et al. (2013) and Richler et al. (2011) reported a correlation between face recognition and an alternative composite effect (see Introduction for differences in the phenomena measured by the standard versus alternative composite effects).

Levine & Calvanio, 1989; Palermo et al., 2011; Ramon et al., 2010; Tanzer, Freud, Ganel, & Avidan, 2013). There are two possibilities. The first possibility is that holistic processing is indeed a hallmark of normal face mechanisms, but variations within normal range are related to non-face processes (e.g. ability to restrict one's attention to the top half of an image). The second possibility is that mechanisms underlying face holistic processing are not needed for successful face identification. However, they may share adjacent brain areas, leading to the observed concurrent damage in many reported cases of acquired prosopagnosia.

We favor the first possibility. The composite and part-whole effect seem to reflect the operation of face-specific mechanisms, but not the efficiency of these mechanisms. How can this be? We can draw an analogy with other visual illusions, such as the Müller-Lyer illusion (Müller-Lyer, 1889). This illusion illustrates that perception of a line length depends on the orientation of the arrows at both ends, but does not say anything about how efficient a visual system is. In other words, it reveals the *type* of information entering computation, or maybe a *method* for computation, but not how efficient the computations are. Similarly, the composite and part-whole effect seem to reflect a type of information (or method) that is critical for successful face recognition (if that information or method are not available, computations cannot proceed), but not how efficiently the face mechanisms process that information.

The null correlation between the composite effect and face recognition is consistent with most previous studies except Wang et al. (2012), and in that study the significant correlation between the two measures was very small (r = .13). There are also a few aspects of the Wang et al. (2012) study that may explain the different results. First, the authors computed the composite effect using both "same" and "different" trials, whereas we, following standard practice (Rossion, 2013), only used "same" trials (no perceptual illusion is expected in the "different" trials). Second, the correlation was obtained only when they used the difference score between face and flower recognition, not when they used the absolute score for face recognition. Third, the presentation times were different from our study (800ms for the first stimulus and 500ms for the second stimulus, compared to 200ms and 200ms in our study). Presentation times longer than 150-200ms allow for multiple saccades and thus may increase the probability of a feature-matching strategy from participants.

4.4. Is the composite effect simply an invalid measure?

The composite effect in our study had essentially null correlations with the inversion effect, the part-whole effect and CFPT. This raises the possibility that there may be something peculiar (methodologically and/or conceptually) about the composite test we used, a possibility that would preclude us from drawing strong inferences from the current data.

However, we believe this is not the case. First, the face composite test produced group effects very similar to those reported in previous studies (Busigny et al., 2010; Jiang, Blanz, & Rossion, 2011; Rossion, 2013; Susilo et al., 2013). Second, the composite test had good internal reliability (higher than the internal reliabilities of the other tests we used), making it an appropriate measure to use in individual differences studies. Third, the composite illusion seems the most straightforward way to capture at least one interpretation of holistic processing, according to which parts *cannot* be represented independently of other parts (Rossion, 2008). The composite effect is a direct measure of the differences in the representation of the *same* face part when other face parts change. In contrast, the part-whole effect contrasts the representations of a face part in *isolation* versus in the face context, while the inversion effect does not tell us much about what is measured.

However, the composite test may be criticized on two fronts. Methodologically, the composite effect is prone to ceiling effects, especially in the misaligned condition. In this case, a difference in performance between the misaligned and aligned trials would reveal the *existence* of perceptual mechanisms engaged solely by aligned faces, but would not be a sensitive quantitative *measure* of the effects of these mechanisms. For example, good top-face recognizers may show less of a composite effect because their at-ceiling scores in the misaligned condition do not reflect their true abilities. This reduced sensitivity to individual differences in the misaligned condition may in principle affect the (lack of) correlations with the other measures. However, that was not the case in our study. Excluding the participants who were at ceiling in the misaligned condition did not modify our findings. Similarly, considering the response times instead of accuracy scores (an option used by some researchers to avoid ceiling effects) produced very similar results.

A group of researchers (e.g. Richler et al., 2011a, 2011b, 2011c, 2011d) has argued that the standard composite design is invalid because it fails to correct for "response bias". In their view, the standard composite effect is driven by a change in decision bias in the aligned versus misaligned condition, with participants more likely to select the "different" response in the aligned condition. However, this group of researchers failed to provide a conceivable mechanism to explain the relevant results: how can a decision bias manifest itself only with upright aligned faces but not with upright misaligned or with inverted aligned and misaligned faces? In contrast, the standard composite design is based on a clear theoretical account to explain the results in all conditions: the upright faces engage *perceptual* mechanisms that are not engaged by stimuli that do not fit a standard face template (i.e. inverted or misaligned faces). Any non-perceptual component in the standard composite effect is controlled for by the

use of inverted and misaligned conditions. In addition, the composite effect in the standard design can also be computed based on response times (instead of the same/different responses) in the aligned versus misaligned "same" conditions. This measure is free of response bias criticisms. In our study, the composite effect based on response times to correct trials produced the same findings (i.e. lack of correlations with the other measures) as the composite effect based on same/different responses. For these reasons, we find the criticism about response bias in the standard composite design does not affect our findings. We also note that the suggested alternative to the standard composite test measures congruency effects which are not face-specific (as opposed to a face-based visual illusion) and so would not have been appropriate for this study (see Introduction for details).

Conceptually, the expectation that the composite effect is a sensitive measure of the efficiency of face-specific mechanisms leads to a paradox. On one hand, more efficient face mechanisms translate into superior face recognition. On the other hand, the size of the composite effect depends mostly on the number of errors in the aligned condition (more errors, larger effect). However, it seems unintuitive to expect *better* face recognizers to perform *worse* (i.e. large composite effect) in a face perception test, regardless of the test. If we take an ideal (i.e. flawless) face recognizer, would we expect her to make the most errors in the aligned condition? Probably not, and one could reasonably argue the opposite. Instead, this observation is consistent with the interpretation of the composite illusion as a hallmark of face-specific information processing, but not as a measure of the efficiency of face perception mechanisms.

4.5. Final considerations

In this paper we considered the common-held belief that the inversion, composite and part-whole effects reflect in one way or another holistic processing. As we have seen, there is compelling evidence to reject this simplistic view. Perhaps the fundamental problem is in trying to measure holistic processing before having a coherent theoretical definition of holistic processing. Despite a few attempts to define it (Galton, 1883; Maurer, Grand, & Mondloch, 2002; Rossion, 2008, 2009; Tanaka & Farah, 1993), it is not clear what face holistic processing is, how it can be computationally implemented (but see Xu, Biederman, & Shah, 2014), or how it is different from the Gestalt-like definition of the emergent properties of a whole (Pomerantz & Portillo, 2011). Future research should aim to more precisely answer these questions. Until then, and until compelling evidence that any of the three measures used in this study measures holistic processing, we advocate the use of the name of each effect (i.e. composite, part-whole, inversion) over the use of the term "holistic processing".

Another question is whether the various mechanisms leading to the three effects, as we have seen mostly at play for upright faces, contribute to face detection. It is conceivable that the ability to instantly integrate information from multiple face parts and produce a whole representation "bigger than the sum of parts" is particularly important for face detection. Admittedly, this would be at odds with evidence from neuropsychology: acquired prosopagnosics typically do not have difficulties with face detection (Bruyer et al., 1983; de Gelder & Rouw, 2000; Rossion et al., 2003; Schiltz et al., 2006) despite reported cases of absent or reduced inversion (Boutsen & Humphreys, 2002; Busigny & Rossion, 2009; McNeil & Warrington, 1991; Rezlescu et al., 2012), reduced composite and/or part-whole effects (Busigny et al., 2010; Delvenne, Seron, Coyette, & Rossion, 2004; Ramon et al., 2010). However, creating sensitive face detection tests is challenging, so subtle impairments may have been overlooked due to ceiling effects. In fact, with rigorous psychophysical testing, a large number of developmental prosopagnosics showed difficulties on face detection tasks (Garrido, Duchaine, & Nakayama, 2008) and a first case of impaired face detection in acquired prosopagnosia has recently been reported (Xu & Biederman, 2014).

Finally, another thing to consider is whether the inversion, composite and part-whole effects tap mechanisms specific to upright faces. There is considerable evidence that inversion produces a disproportionately larger effect in faces than non-face objects, but a small inversion effect is sometimes found with non-face objects. The evidence is more mixed when it comes to the composite and part-whole effects. The reduced (or lack of) effects with non-face objects may be due to stimuli confounds. Most non-face stimuli are rigid objects that can be reliably identified by parts. Their parts typically do not change shape and have constant spatial relations among them, therefore identification of a single difference between two objects is a reliable indication they are different. When non-face stimuli are parametrically manipulated (e.g. Busigny et al., 2010), the changes follow predictable paths, the stimuli are clearly artificial and the smallest detectable difference leads again to perception of two distinct objects because of expected rigidness. For biological stimuli (e.g. bodies) that have non-rigid parts with varying distances between them, the scale of the differences between two objects is usually much higher than for faces. Definitive conclusions about the specificity of presumed face-specific behavioral effects can be drawn only by comparison of faces to natural, non-face looking objects matched to faces in terms of scope, magnitude and plausibility of part changes.

If the inversion, composite and part-whole effects are not specific to upright faces, they may reflect more general Gestalt processing mechanisms, of the type measured in visual closure tests (Street, 1931) or in the Navon task (Navon, 1977). Even if that was the case, it is highly

unlikely there is a single set of mechanisms that can be defined as holistic. Note that a recent study (Gerlach & Krumborg, 2014) found that even the Navon task may tap different Gestalt mechanisms depending on the stimulus type (letters versus shapes). Therefore, to avoid any ambiguities regarding the (lack of) unity of holistic mechanisms, a more accurate (and acceptable) description of the mechanisms reflected by the inversion, composite, part-whole and Navon effects (as well as others) is that they are of the holistic *type* (in a very broad sense).

5. Conclusion

Discussing holistic processing as an all-encompassing term is misleading. The inversion, composite and part-whole effects tap distinct mechanisms and only the inversion effect could reasonably predict face recognition abilities. The mechanisms behind the inversion, composite and part-whole effects may be described as of a holistic type, but the distinctions among them should be emphasized in studies aimed at understanding holistic processing and the nature of face perception mechanisms.

Table 1. Reliabilities and overall performance for the four tests, per web sample, lab sample, and for all participants. The results for each condition in the tests are in bold, while the computed effects are in normal fonts. The second column presents the formulas used for computation, e.g. A-B = the difference after subtracting B from A; A~B = the residuals after regressing out B from A.

		RELIABILITY Guttman's λ2			OVERALL PERFORMANCE Means			SD		
		web	lab	all	web	lab	all	web	lab	all
Cambridge Face Perception Test										
Upright	**A**	**0.67**	**0.71**	**0.68**	**0.72**	**0.74**	**0.72**	**0.10**	**0.10**	**0.10**
Inverted	**B**	**0.50**	**0.60**	**0.52**	**0.53**	**0.54**	**0.53**	**0.10**	**0.11**	**0.10**
Inversion effect subtraction	A-B	0.15	0.30	0.18	0.19	0.20	0.19	0.10	0.10	0.10
Inversion effect regression	A~B	0.37	0.47	0.39	-	-	-	-	-	-
Face Matching Test										
Upright	**C**	**0.77**	**0.78**	**0.77**	**0.81**	**0.82**	**0.82**	**0.10**	**0.10**	**0.10**
Inverted	**D**	**0.57**	**0.60**	**0.57**	**0.54**	**0.55**	**0.54**	**0.09**	**0.09**	**0.09**
Inversion effect subtraction	C-D	0.35	0.29	0.32	0.27	0.28	0.27	0.10	0.09	0.09
Inversion effect regression	C~D	0.56	0.49	0.53	-	-	-	-	-	-
Face Composite Test										
Upright aligned	**E**	**0.80**	**0.80**	**0.80**	**0.62**	**0.64**	**0.63**	**0.18**	**0.18**	**0.18**
Upright misaligned	**F**	**0.66**	**0.70**	**0.67**	**0.89**	**0.88**	**0.89**	**0.09**	**0.10**	**0.09**
Upright composite effect subtraction	F-E	0.61	0.66	0.62	0.27	0.24	0.26	0.17	0.18	0.18
Upright composite effect regression	E~F	0.74	0.77	0.75	-	-	-	-	-	-
Inverted aligned	**G**	**0.74**	**0.76**	**0.75**	**0.84**	**0.81**	**0.83**	**0.13**	**0.13**	**0.13**
Inverted misaligned	**H**	**0.68**	**0.78**	**0.72**	**0.89**	**0.87**	**0.89**	**0.09**	**0.12**	**0.10**
Inverted composite effect subtraction	H-G	0.40	0.15	0.35	0.06	0.06	0.06	0.11	0.09	0.11
Inverted composite effect regression	G~H	0.53	0.24	0.46	-	-	-	-	-	-
Upright-specific composite effect subtraction	(F-E)-(H-G)	0.34	0.22	0.28	0.22	0.18	0.21	0.18	0.17	0.18
Upright-specific composite effect regression	(E~F)~(G~H)	0.68	0.68	0.66	-	-	-	-	-	-
Part-Whole Test										
Upright whole trials	**I**	**0.82**	**0.81**	**0.81**	**0.78**	**0.75**	**0.77**	**0.10**	**0.10**	**0.10**
Upright part trials	**J**	**0.63**	**0.63**	**0.62**	**0.69**	**0.68**	**0.69**	**0.08**	**0.08**	**0.08**
Upright part-whole effect subtraction	I-J	0.28	0.23	0.25	0.09	0.07	0.08	0.08	0.08	0.08
Upright part-whole effect regression	I~J	0.48	0.43	0.45	-	-	-	-	-	-
Inverted whole	**K**	**0.60**	**0.58**	**0.59**	**0.65**	**0.62**	**0.64**	**0.08**	**0.08**	**0.08**
Inverted part	**L**	**0.48**	**0.29**	**0.42**	**0.62**	**0.61**	**0.62**	**0.07**	**0.06**	**0.07**
Inverted part-whole effect subtraction	K-L	0.23	0.18	0.21	0.03	0.01	0.03	0.09	0.09	0.09
Inverted part-whole effect regression	K~L	0.42	0.46	0.43	-	-	-	-	-	-
Upright-specific part-whole effect subtraction	(I-J)-(K-L)	0.03	-0.04	0.04	0.05	0.06	0.05	0.11	0.11	0.11
Upright-specific part-whole effect regression	(I~J)~(K~L)	0.38	0.32	0.37	-	-	-	-	-	-

Figure 1. Correlations between the three measures of holistic processing. The grey bars represent the maximum possible correlation considering test reliabilities. Significant correlations (p<.05) are in green, nonsignificant correlations are in red. The triangle in the middle shows the significant differences between correlations. n.s. = not significant.
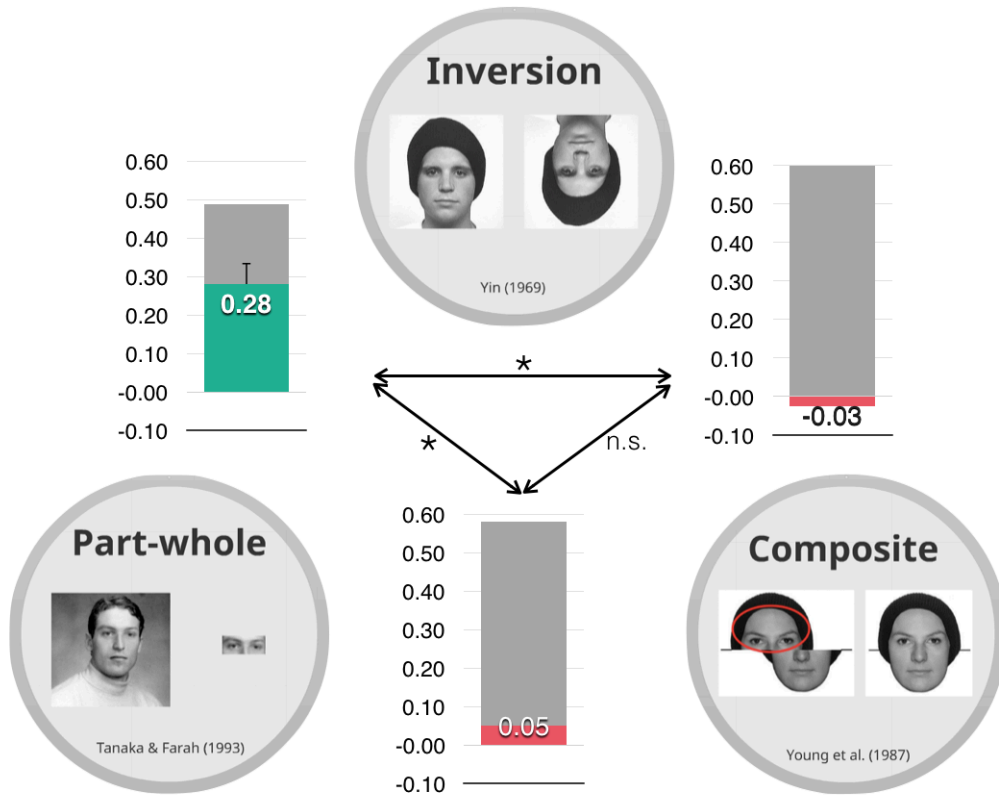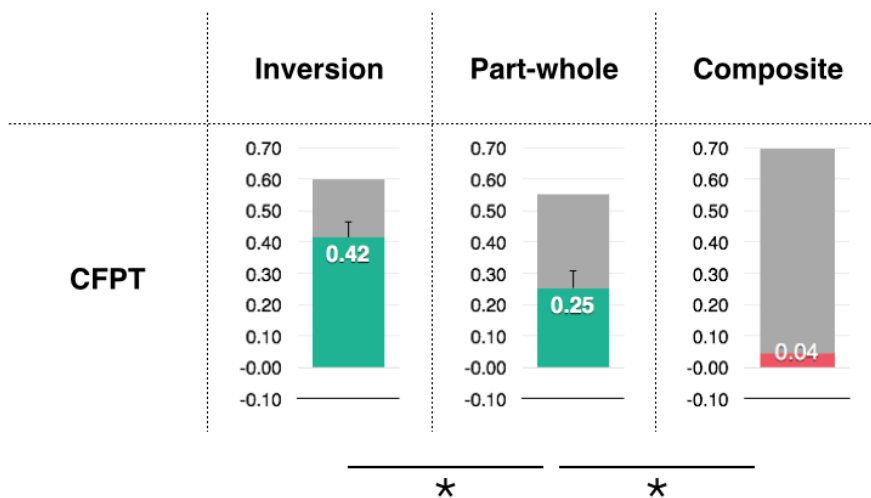

Figure 2. Correlations between measures of holistic processing (three columns) and face perception abilities (CFPT). The grey bars represent the maximum possible correlations considering test reliabilities. Significant correlations (p<.05) are in green, nonsignificant correlations in red.

Acknowledgments

Funding

References

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, *49*(9), 2541–52. doi:10.1016/j.neuropsychologia.2011.05.002

Behrmann, M., Richler, J. J., Avidan, G., & Kimchi, R. (in press). Holistic face perception. In J. Wagemans (Ed.), *The Oxford handbook of perceptual organization*. Oxford University Press.

Boutsen, L., & Humphreys, G. W. (2002). Face context interferes with local part processing in a prosopagnosic patient. *Neuropsychologia*, *40*(13), 2305–13.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., … Yovel, G. (2009). Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, *26*(5), 423–55. doi:10.1080/02643290903343149

Bruyer, R., Laterre, C., Seron, X., Feyereisen, P., Strypstein, E., Pierrard, E., & Rectem, D. (1983). A Case of Prosopagnosia with Some Preserved Covert Remembrance of Familiar Faces. *Brain and Cognition*, *2*, 257–284.

Bukach, C. M., Phillips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, and Psychophysics*, 72, 1865-1874. doi:10.3758/APP.72.7.1865

Busigny, T., Joubert, S., Felician, O., Ceccaldi, M., & Rossion, B. (2010). Holistic perception of the individual face is specific and necessary: Evidence from an extensive case study of acquired prosopagnosia. *Neuropsychologia*, 48(14), 4057–4092. doi:10.1016/j.neuropsychologia.2010.09.017

Busigny, T., & Rossion, B. (2009). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, *46*(8), 965–981. doi:10.1016/j.cortex.2009.07.004

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology. Human Perception and Performance*, 34(6), 1327–36. doi:10.1037/a0011752

de Gelder, B., & Rouw, R. (2000). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? *Neuroreport*, *11*(14), 3145–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11043539

DeGutis, J., Cohan, S., Mercado, R. J., Wilmer, J., & Nakayama, K. (2012). Holistic processing of the mouth but not the eyes in developmental prosopagnosia. *Cognitive Neuropsychology*, *29*(5-6), 419–46. http://doi.org/10.1080/02643294.2012.754745

DeGutis, J. M., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, 126(1), 87–100. doi:10.1016/j.cognition.2012.09.004

Delvenne, J.-F., Seron, X., Coyette, F., & Rossion, B. (2004). Evidence for perceptual deficits in associative visual (prosop)agnosia: a single-case study. *Neuropsychologia*, *42*(5), 597–612. doi:10.1016/j.neuropsychologia.2003.10.008

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective neurons in the macaque. *Journal of Neuroscience*, *4*(8), 2051–2062.

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology*, 24, 419–430. doi:10.1080/02643290701380491

Duchaine, B., & Yovel, G. (2008). Face recognition. In A. Basbaum, R. Hoy, A. Kaneko, G. Shepherd, & G. Westheimer (Eds.), *The Senses: A Comprehensive Reference* (Vol. 2, pp. 329–358). Amsterdam: Elsevier.

Duchaine, B., Yovel, G., & Nakayama, K. (2007). No global processing deficit in the Navon task in 14 developmental prosopagnosics. *Social Cognitive and Affective Neuroscience*, 2, 104–113.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. W. (1998). What is "special" about face perception? *Psychological Review*, *105*(3), 482–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9697428

Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience*, 12(9), 1187–1196. doi:10.1038/nn.2363

Galton, F. (1883). Inquries into human faculty and its development. London, UK: Macmillan

Garrido, L., Duchaine, B., & Nakayama, K. (2008). Face detection in normal and prosopagnosic individuals. *Journal of Neuropsychology*, *2*(1), 119–140. http://doi.org/10.1348/174866407X246843

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, 6, 428?432. doi:10.1038/nn1029

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behaviour. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 431-466. doi:10.1037/0096-1523.28.2.431

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. W. (1998). *Training ''greeble'' experts: A framework for studying expert object recognition processes*. Vision Research, 38, 2401-2428. doi:10.1016/S0042-6989(97)00442-2

Gerlach, C., & Krumborg, J. R. (2014). Same, same - but different: On the use of Navon derived measures of global/local processing in studies of face processing. *Acta psychologica*, 153C, 28–38. doi:10.1016/j.actpsy.2014.09.004

Gignac, G. E., Shankaralingam, M., Walker, K., & Kilpatrick, P. (2016). Short-term memory for faces relates to general intelligence moderately. *Intelligence*, *57*, 96–104.

Green, C. D. (1992). Of immortal mythological beasts: operationism in psychology. *Theory & Psychology*, 2, 291–320. doi: 10.1177/0959354392023003

Hayward, W. G., Crookes, K., & Rhodes, G. (2013). The other-race effect: Holistic coding differences and beyond. *Visual Cognition*, *21*(9-10), 1224-1247.

Hobson, R. P., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British Journal of Psychology*, *79*(4), 441-453.

Jiang, F., Blanz, V., & Rossion, B. (2011). Holistic processing of shape cues in face identification: Evidence from face inversion, composite faces, and acquired prosopagnosia. *Visual Cognition*, *19*(8), 1003–1034. doi:10.1080/13506285.2011.604360

Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry*, *44*(4), 529-542.

Kanwisher, N. G., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science*, 21(1), 38–43. doi:10.1177/0956797609356508

Leahey, T. H. (1980). The myth of operationism. *Journal of Mind and Behavior*, Vol 1(2), 1980, 127-143.

Lee, I. A., & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org.

Levine, D. N., & Calvanio, R. (1989). Prosopagnosia: A defect in visual configural processing. *Brain and Cognition*, 10, 149–170.

Lilienfeld, S. O., Sauvigné, K. C., Lynn, S. J., Cautin, R. L., Latzman, R. D., & Waldman, I. D. (2015). Fifty psychological and psychiatric terms to avoid: a list of inaccurate, misleading, misused, ambiguous, and logically confused words and phrases. *Frontiers in Psychology*, 6(August), 1–15. doi:10.3389/fpsyg.2015.01100

Liu, J., Harris, A., & Kanwisher, N. G. (2010). Perception of face parts and face configurations: an FMRI study. *Journal of Cognitive Neuroscience*, 22(1), 203–11. doi:10.1162/jocn.2009.21203

Maurer, D., Grand, R. Le, & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. doi:10.1016/S1364-6613(02)01903-4

McKone, E., Crookes, K., & Jeffery, L., & Dilks, D.D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology*, 29(1), 174-212.

McKone, E., Kanwisher, N. G., & Duchaine, B. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), 8–15. doi:10.1016/j.tics.2006.11.002

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.

McNeil, J. E., & Warrington, E. K. (1991). Prosopagnosia: a reclassification. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(2), 267–87. doi:10.1080/14640749108400970

Michel, C., Rossion, B., Han, J., Chung, C.-S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of our own race. *Psychological Science*, 17, 608–615.

Müller-Lyer, F.C. (1889). Optische Urteilstäuschungen. *Archiv für Physiologie Suppl.* 263–270.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, 49(5), 1226–35. doi:10.1016/j.neuropsychologia.2011.02.021

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science*, 14(6), 618–622.

Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms "holistic", "configural," and "relational" in the face perception literature. *Frontiers in Psychology*, 3 (559).

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1331–1349. doi:10.1037/a0024330

Ramon, M., Busigny, T., & Rossion, B. (2010). Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. *Neuropsychologia*, *48*(4), 933–44. doi:10.1016/j.neuropsychologia.2009.11.014

Rezlescu, C., Pitcher, D., & Duchaine, B. (2012). Acquired prosopagnosia with spared within- class object recognition but impaired recognition of degraded basic-level objects. *Cognitive Neuropsychology*, *29*(4), 325–347.

Rhodes, G. (2013). Face recognition. In D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology* (pp. 46-68). New York: Oxford University Press.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011a). Holistic processing predicts face recognition. *Psychological Science*, 22(4), 464–71. doi:10.2307/330093

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011b). Beliefs alter holistic face processing … if response bias is not taken into account. *Journal of Vision*, 11(13), 1–13.

Richler, J. J., Floyd, R. J., & Gauthier, I. (2014). The Vanderbilt Holistic Face Processing Test: A short and reliable measure of holistic face processing. *Journal of Vision*, *14*(11), 1–14.

Richler, J. J., Floyd, R. J., & Gauthier, I. (2015). About-face on face recognition ability and holistic processing. *Journal of Vision*, *15*(9), 15.

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

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic processing. *Psychological Bulletin*, 140(5), 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(2), 328–342. doi:10.1037/0278-7393.34.2.328

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011c). Inverted faces are (eventually) processed holistically. *Vision Research*, 51(3), 333–42.

Richler, J. J., Wong, Y. K., & Gauthier, I. (2011d). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science*, 20(2), 129–134. doi:10.1177/0963721411402472

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

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: the perceptual field hypothesis. *Acta Psychologica*, 132, 300-312.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139–253. doi:10.1080/13506285.2013.772929

Rossion, B., Caldara, R., Seghier, M. L., Schuller, A.-M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, *126*(Pt 11), 2381–95. doi:10.1093/brain/awg241

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*(3), 1385–1394.

Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., & Rossion, B. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cerebral Cortex*, *16*(4), 574–86. doi:10.1093/cercor/bhj005

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*(5), 391–396. http://doi.org/10.1016/j.cub.2004.02.028

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245-251.

Street, R. F. (1931). A Gestalt completion university. New York, NY: Teachers College of Columbia University.

Susilo, T., Crookes, K., McKone, E., & Turner, H. (2009). The composite task reveals stronger holistic processing in children than adults for child faces. *PLoS One*, *4*(7), e6460.

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Pidcock, M., et al. (2010). Face recognition impairments despite normal holistic processing and face space coding: evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology*, 27(8), 636–64. doi:10.1080/02643294.2011.613372

Susilo, T., Rezlescu, C., & Duchaine, B. (2013). The composite effect for inverted faces is reliable at large sample sizes and requires the basic face configuration. *Journal of Vision*, *13*(13), 1–9. doi:10.1167/13.13.14.doi

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, 46(2), 225–245. doi:10.1080/14640749308401045

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

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition*, *93*(1), B1–B9. doi:10.1016/j.cognition.2003.09.011

Tanzer, M., Freud, E., Ganel, T., & Avidan, G. (2013). General holistic impairment in congenital prosopagnosia: evidence from Garner's speeded-classification task. *Cognitive Neuropsychology*, 30(6), 429–45. doi:10.1080/02643294.2013.873715

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science*, 23(2), 169–77. doi:10.1177/0956797611420575

Wong, A. C.-N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. G. H. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision*, 12(7), 1-15. doi:10.1167/12.7.1

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

Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PLoS ONE*, 6(6), e20753. doi:10.1371/journal.pone.0020753

Xu, X., & Biederman, I. (2014). Neural Correlates of Face Detection. *Cerebral Cortex*, *24*(June), 1555–1564. doi:10.1093/cercor/bht005

Xu, X., Biederman, I., & Shah, M. P. (2014). A neurocomputational account of the face configural effect. *Journal of Vision*, *14*(8), 1–9. http://doi.org/10.1167/14.8.9

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

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747–759. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3454432

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., … Liu, J. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology*, *20*(2), 137–42. doi:10.1016/j.cub.2009.11.067