A Meta-Analysis and Review of Holistic Face Processing

Jennifer J. Richler and **Isabel Gauthier**
Department of Psychology, Vanderbilt University

Abstract

The concept of holistic processing is a cornerstone of face recognition research, yet central questions related to holistic processing remain unanswered, and debates have thus far failed to reach a resolution despite accumulating empirical evidence. We argue that a considerable source of confusion in this literature stems from a methodological problem. Specifically, two different measures of holistic processing based on the composite paradigm (complete design and partial design) are used in the literature, but they often lead to qualitatively different results. First, we present a comprehensive review of the work that directly compares the two designs, and which clearly favors the complete design over the partial design. Second, we report a meta-analysis of holistic face processing according to both designs, and use this as further evidence for one design over the other. The meta-analysis effect size of holistic processing in the complete design is nearly three times that of the partial design. Effect sizes were not correlated between measures, consistent with the suggestion that they do not measure the same thing. Our meta-analysis also examines the correlation between conditions in the complete design of the composite task, and suggests that in an individual differences context, little is gained by including a misaligned baseline. Finally, we offer a comprehensive review of the state of knowledge about holistic processing based on evidence gathered from the measure we favor based on the first sections of our review—the complete design—and outline outstanding research questions in that new context.

Keywords

face perception; holistic processing; individual differences; expertise

A visual preference for faces is observed during infancy, and infants spend more time looking at or tracking faces compared with other highly salient visual stimuli (e.g., Johnson, et al., 1991; Maurer & Young, 1983). Moreover, despite the limitations of the newborn visual system, infants are able to discriminate their mother's face from the face of a stranger solely on the basis of visual information (e.g., Bushnell et al., 1989). Faces are important to our social interactions, are omnipresent in our environment from infancy, and are associated with speech and other salient sensory information. Early prioritization of face stimuli could be due to any combination of these factors. This lays the seeds for a very important debate in the literature on face perception: are faces inherently "special", such that there is an innate genetic predisposition to attend to face-like stimuli (e.g., Morton & Johnson, 1991), or are processing differences between faces and other objects driven by differences in amount or

type of experience, with infants rapidly learning about faces because they are dominant in their early environment (e.g., Butko, Fasel & Movellan, 2006)?

Whether specialization for faces is the result of an innate face module (McKone, Kanwisher & Duchaine, 2007) or can be accounted for by experience with faces (Bukach, Gauthier & Tarr, 2006) is central to understanding the organization of brain areas involved in visual object processing and how such organization gives rise to behavior. In particular, the lateral fusiform gyrus shows such strong selectivity to faces compared to other objects, including scrambled faces, that it has been named the Fusiform Face Area (FFA; Kanwisher, McDermott & Chun, 1997; see Weiner & Grill-Spector, 2013 for a review). However, this area has also been shown to be selective for non-face objects in individuals who are considered experts with those objects (e.g., birds in bird watchers; Gauthier et al., 2000; Xu, 2005; Engel et al., 2009; McGugin et al., 2012a).

Moreover, whether or not face perception is distinct from the recognition of other objects is relevant to interpreting recent claims that face recognition is a reliable ability (e.g., Duchaine & Nakayama, 2006) that is highly heritable (Wilmer et al., 2012; Zhu et al., 2010). The heritability of performance on face recognition tests is consistent with evidence for a hereditary form of a face recognition deficit, prosopagnosia (e.g., De Haan, 1999; Grueter, Grueter & Carbon, 2008). But a more critical question is what is this ability that underlies performance on face recognition tests? Understanding the processes and behaviors that are specific to face recognition is critical to understanding the nature of what exactly is being inherited.

Holistic Processing

Many different behavioral tasks and measures have been used to test the claim that faces are processed differently than other objects. One of the most prominent ideas is that faces are processed holistically— as unified wholes or "gestalts" — compared to other objects that are processed in a more analytic, part- or feature-based manner (Farah et al., 1998). Unlike objects that are typically identified at the category-level (e.g., "dog"; Rosch et al., 1976), it is often necessary to identify faces at the level of the individual (e.g., "Bob"). But, all faces consist of the same kinds of features (eyes, nose, and mouth) in the same configuration (eyes above nose, nose above mouth). Holistic processing is believed to facilitate individuation of such visually similar objects.

Holistic processing of faces was initially explained by a holistic encoding or template hypothesis, according to which faces are encoded as a single unit to fit a "face template", making it difficult to operate on face parts that are not explicitly represented (Farah et al., 1998; Tanaka & Farah, 1993). The idea that holistic processing occurs during encoding is consistent with several different computer vision algorithms of face detection. For example, because of the high consistency in the arrangement of features, and by extension dark and light patches in all faces, the presence of a face can be ruled out rapidly by very simple filters or overlapping elemental feature detectors. The feature overlaps in these models implicitly enforce the correct overall arrangement of features, and holistic processing could

be the result of this detection stage that uses a coarse upright template (see Tsao & Livingstone, 2008, for a review).

An alternative to the template hypothesis is that holistic processing is the outcome of an attentional strategy that has become automatized with experience (Richler, Palmeri, & Gauthier, 2012; Richler, Wong, & Gauthier, 2011). According to this view, face parts are encoded and represented independently, and holistic processing reflects arises from obligatory encoding of all object parts because a strategy of attending to all parts cannot be "turned off". A related view is that independently represented face parts are not treated independently during perceptual decision-making, which could be the cause of holistic effects (e.g., Richler, Gauthier, Wenger & Palmeri, 2008; Richler, Tanaka, Brown & Gauthier, 2008).

Many measures have been used to assess holistic processing, but different tasks and measures are associated with different ideas of what holistic processing might mean. For example, the composite task (Young et al., 1987) indexes the extent to which participants can ignore irrelevant information in a face — treating faces as wholes means that participants cannot selectively attend to face parts, even when it is disadvantageous for the task. In contrast, the part-whole task (Tanaka & Farah, 1993) measures an advantage for the whole over individual parts — treating faces as wholes means that the whole is greater than the sum of its parts. We have described this issue in detail elsewhere (Richler, Palmeri, & Gauthier, 2012), but for the present purpose it is important to emphasize that the results from different measures of face processing need not converge, because it is likely that many different processes contribute to face perception, and they do not have to be equally influenced by all possible manipulations. Similarly, although the term "configural processing" is often used interchangeably with "holistic processing" (e.g., Crookes & McKone, 2009; DeGutis et al., 2011; McKone, 2008), these are at least in principle constructs that can be dissociated. Configural processing refers to sensitivity to spatial relations between parts (e.g., distance between the eyes); holistic processing is the tendency to process all parts together. Holistic processing may facilitate configural processing, but these two processing strategies that are beneficial in face perception may be independent, differentially affected by experimental manipulations (see Maurer, Le Grand & Mondloch, 2002 for a review), and have different developmental trajectories (e.g., Carey & Diamond, 1994, vs. Freire & Lee, 2001; Mondloch et al., 2007).

To preview one contentious example, it is well-known that inversion disrupts discrimination and memory for faces more than for other objects (e.g., Yin, 1969). But this on its own does not imply that inversion must disrupt holistic processing. Inversion might disrupt configural and/or feature processing (see McKone & Yovel, 2009, for a review), or reduce overall processing efficiency (Curby & Gauthier, 2009; Sekuler et al., 2004) because of less effective fixation patterns (Hills, Cooper & Pake, 2013; Xu & Tanaka, 2013). In other words, stronger inversion effects for faces vs. objects can be accounted for by other factors besides holistic processing, and holistic processing is not in any principled way necessary to explain the inversion effect.

Although many reviews lump various tasks together to draw general conclusions about holistic processing (e.g., McKone, Kanwisher & Duchaine, 2007), we do not believe that this is a prudent approach given the poor evidence that these measures map on to the same construct (see Richler, Palmeri, & Gauthier, 2012). While many studies show group effects for different tasks in the same direction (e.g, McKone, 2009), there is much less work looking at correlations across tasks, and even when these relationships are found, the magnitude of the correlations is modest (e.g., DeGutis et al., 2013). It should be noted, however, that most of these tasks were not created for psychometric purposes and the reliability of the measurements needs to be improved (see Ross, Richler, & Gauthier, under review). Therefore, here we will focus on what is arguably the most popular task used by the field to measure holistic processing, the composite task. Note that we believe that, in principle, using multiple indicators is generally preferable. However, for this one needs several valid indicators of the same construct; it is our hope that with more attention to measurement in the field of face recognition, this approach will eventually be viable.

The Composite Task

In a landmark study by Young and colleagues (1987), naming latencies for the top half of a famous face were slower when it was aligned with the bottom half of a face belonging to a different famous individual compared to when the two face halves were misaligned. The authors called this effect an "illusion" because when aligned, the two famous face halves form a new unfamiliar face that is so convincing it interferes with identification of the constituent parts. The same effect is obtained in matching paradigms with unfamiliar faces. It is more difficult to judge whether the top halves of two faces are the same or different when they are aligned with different bottom halves compared to when the face halves are misaligned (e.g., Hole, 1994).

The composite effect indexes failures of selective attention: participants cannot ignore the information in the irrelevant face half. When the face halves are misaligned, the meaningful face configuration is disrupted, reducing or eliminating holistic processing (e.g., Hole, 1994; Richler, Tanaka, Brown & Gauthier, 2008). Note that the phrase "failure of selective attention" is an operationalization of the task — a to-be-ignored part influences judgments about a target part — and should not be interpreted as a statement about the underlying mechanism (see Richler & Gauthier, 2013; Richler, Palmeri, & Gauthier, 2012). In fact, the mechanism that gives rise to these failures of selective attention is disputed (see Richler & Gauthier, 2013; Rossion, 2013).

Importantly, holistic processing in the composite task is not observed for non-face objects in novices (Farah et al., 1998; Richler, Mack et al., 2011), suggesting that it is face-specific. But, holistic processing is observed for non-face objects in real-world experts (e.g., cars in car experts; Bukach et al., 2010) and for novel objects following individuation training (Gauthier et al., 1998; Wong et al., 2009a). Moreover, the magnitude of holistic processing for non-face objects is correlated with FFA activity for those objects following individuation training (Gauthier & Tarr, 2002; Wong et al., 2009b). Although these results support an expertise hypothesis, where faces are processed differently from other objects because they are a category with which we are experts, these results can be difficult to interpret because

others do not find holistic processing for non-face objects of expertise in the first place (Robbins & McKone, 2007).

Beyond its relevance to the expertise debate, there are indications that holistic processing has gained a central role in the broader literature on face processing, and the composite task in particular has reached a stage where it is considered somewhat of a standard paradigm in related areas of research. The composite task has been used to track the development of face recognition (e.g., Cassia et al., 2009; Mondloch et al., 2007), to study abnormal development of face recognition (e.g., developmental prosopagnosics; Le Grand et al., 2006) and populations with face recognition deficits as part of more widespread cognitive impairment (e.g., schizophrenia; Schwartz et al., 2002), and to evaluate computational models of face recognition (Dailey & Cottrell, 1999). On the one hand, a task that is adopted as a standard measure and the widespread investigation of holistic processing are important contributions from the face processing literature to other domains of research. On the other hand, the value of the contribution is constrained by the validity of the tool from both measurement and theoretical perspectives. Indeed, disagreements about measurement are not simply technical —they are important to understand for anyone faced with the choice of which measure to use for a construct of interest. Indeed, although some studies find that holistic processing is related to face recognition abilities (DeGutis et al., 2013; McGugin, Richler et al., 2012; Richler, Cheung, & Gauthier, 2011b), supporting the widespread use of the composite task as a tool to understand face recognition, others do not (Konar, Bennett & Sekuler, 2010; Wang et al., 2012[1]).

This is not the only example of discrepancies in the literature on holistic processing. For instance, while some studies suggest that holistic processing is mainly supported by low-spatial frequency information (Goffaux & Rossion, 2006; Goffaux, 2009), other work finds that high-pass filtered faces are processed as holistically as low-pass filtered faces (Cheung et al., 2008); according to some faces (Curby, Goldstein, & Blacker, 2013; Richler, Mack, et al., 2011), while other studies find that inversion abolishes holistic processing completely (e.g., Rossion & Boremanse, 2008); as mentioned above, holistic processing has been observed for non-face objects following the acquisition of perceptual expertise (Boggan, Bartlett & Krawczyk, 2012; Gauthier & Tarr, 2002; Wong et al., 2009a), but other studies fail to replicate this result (Robbins & McKone, 2007).

One way to step back from these empirical inconsistencies would be to conduct a meta-analysis, an approach that is rarely used in the face perception literature (see Meissner & Brigham, 2001, and McKone & Yovel, 2009, for exceptions). However, one key principle of meta-analysis is that one should not aggregate apples and oranges. That is, meta-analytic methods allow for a lot of noise in estimation, but not for conceptual variability. Rather than reflecting interesting differences in experimental manipulations, inherent instability of holistic processing, or noise in the measurement of holistic processing, our recent work leads

---

[1]Although these authors found a correlation between holistic processing and performance on a face task when an object baseline was subtracted, the individual correlations with the measure of holistic processing revealed no correlation with face recognition and a small but negative correlation with the control object task (Jia Liu, personal communication). Thus, these results support an association between the measure of holistic processing and object processing, not face processing (see Humphreys, 1990, for a discussion of common errors in the interpretation of correlations with difference scores).

us to conclude that inconsistencies between studies arise from a common source of confusion in this literature, one that stems from a fundamental problem regarding construct validity. There are currently two different versions of the composite task in the literature, and these two versions can be largely blamed for qualitative empirical inconsistencies. This makes it difficult — if not impossible — to integrate the literature on face perception and holistic processing: how do we relate empirical results to each other when differences in holistic processing between studies arise because the measures of holistic processing are qualitatively different? It is not the case that the two versions of the composite task lead to different results and conclusions because they are based on different possible meanings of holistic processing (cf. Richler, Palmeri & Gauthier, 2012). Indeed, from the perspective of the participant, the task is identical in both versions of the composite task.

In the next section, we argue that one of these measures of holistic processing should be abandoned because it confounds manipulations of interest with complex response biases, resulting in poor construct validity. Understanding why the results from one version of this task are not valid is critical to resolving many of the extent debates in the literature about holistic processing that constrain our understanding of face perception; in most cases, these debates are fueled by the fact that the two versions of the composite task yield different results, and they are unlikely to converge unless the measurement issues are resolved. Then, we present the first meta-analysis of holistic processing that further supports one measure over the other, and that provides insight into other important measurement issues associated with holistic processing. Finally, we provide a comprehensive review of holistic processing and face recognition based on studies that use the design we argue is more valid.

A Tale of Two Measures

The two versions of the composite task that are currently used in the literature (called the *partial design* and the *complete design*; Gauthier & Bukach, 2007) are both sequential matching tasks with unfamiliar faces. Participants see one (study) composite face and, following a brief delay, are then presented with a second (test) composite face. Participants are asked to report whether the target half of the test composite (either top or bottom) is the same or different as the corresponding part of the study composite (see Figure 1). Additionally, both versions of the composite task operationalize holistic processing as a failure of selective attention: because faces are processed as wholes it is difficult to selectively attend to the target face half, resulting in interference from the task-irrelevant face half.

The Partial Design

The more popular partial design of the composite task is similar to that of Hole (1994), who first offered a matching version of the naming task used in Young et al. (1987). The target face half is either the same as or different from the corresponding part of the study face, but the irrelevant halves are always different (see Figure 1). Performance on these trials with aligned parts is compared to the same trials where the parts are misaligned. Holistic processing is inferred from an *alignment effect*[2] accuracy is higher on misaligned vs. aligned trials.

Advocates of the partial design have pointed out that predictions in this task can only be made about "same" trials (Robbins & McKone, 2007; Rossion, 2013): when face parts are aligned the different irrelevant parts should hurt performance on "same" trials because they cannot be ignored. In contrast, it is much more difficult to make a prediction about the "different" trials. For example, looking at Figure 1, the different B and D bottoms could increase the perceived difference that already exists between the A and C tops, but it could also reduce the perceived difference depending on whether B is more similar to D than A is to C. For this reason, some authors analyze accuracy on "same" trials only in the partial design (e.g., Michel et al., 2006; Susilo et al., 2013).

The Complete Design

The complete design, first used by Farah, Wilson, Drain & Tanaka (1998), includes the other half of the condition matrix, adding "same" and "different" trials with irrelevant face halves that are the same. Accordingly, one can define "congruent" and "incongruent" trials, depending on the relation between the correct response (same or different) for the target part and the same/different status of the task-irrelevant, to-be-ignored part (see Figure 1). On congruent trials, the response to the target part matches the same/different status of the irrelevant part (i.e., both parts are the same or both parts are different). On incongruent trials, the response to the target part conflicts with the same/different status of the irrelevant part (i.e., one part is the same, one part is different). Holistic processing is often inferred from a *congruency × alignment interaction* (e.g., Richler, Cheung & Gauthier, 2011a,b; Richler, Mack et al., 2011): performance (measured in sensitivity; d') is better on congruent vs. incongruent trials, and the magnitude of this congruency effect is reduced when parts are misaligned.

In some earlier work, a misaligned condition was not included and the congruency effect on aligned trials was used to index holistic processing (e.g., Richler, Gauthier et al., 2008; Richler, Mack et al., 2009). However, later work showed that congruency effects that are not modulated by alignment can sometimes be observed in novices for strategic reasons (e.g., Richler, Bukach & Gauthier, 2009; Wong & Gauthier, 2010). Importantly, the congruency × alignment interaction can be used to check that the effect is not obtained outside of the familiar configuration (see Richler, Wong & Gauthier, 2011, for a review; and Richler & Gauthier, 2013, for a discussion of alternative baselines). We will return to this point in our meta-analysis.

Unlike the partial design, in the complete design predictions about holistic processing can be made for both same and different trials. If participants cannot selectively attend to the target part, they will be relatively impaired on incongruent relative to congruent trials. On same-incongruent trials (i.e., trials where the correct response is "same" but the irrelevant part is different), failure to selectively attend to the target parts can only make participants more likely to respond "different", and on different-incongruent trials (i.e., trials where the correct response is "different" but the relevant part is the same), failures of selective attention can

---

[2]That this measure of holistic processing depends directly on configuration may be at the origin of the suggestion by some that holistic and configural processing are the same thing (Rossion, 2013). Nonetheless, we here consider that the partial design has reasonable face validity as a measure of holistic processing because participants are instructed to ignore a part and fail to do so.

only make participants more likely to respond "same". Because faces are processed as wholes, information from the irrelevant part cannot be ignored and interferes with performance when the information is contradictory. Importantly, although this task is similar to a classic Stroop task (Stroop, 1935), unlike Stroop effects, interference indicative of holistic processing arises at a perceptual level, and does not reflect response interference (Cheung, Richler, Phillips, & Gauthier, 2011; Richler, Cheung, Wong, & Gauthier, 2009).

The Problem with the Partial Design

To understand the problem with the partial design, we need to step back and consider a two-alternative forced-choice task, like a same/different task, in the context of Signal Detection Theory (SDT; Green & Swets, 1966). To foreshadow our argument, it is not simply that SDT should be used to account for response bias; the problem is that the partial design confounds level of congruency with correct response, such that any response bias associated with congruency gets lumped together with measures of holistic processing. Understanding the problem therefore requires a quick overview of how SDT applies to this task.

According to SDT, making a judgment, like a same/different judgment, is the end result of at least two processes: a perceptual process and a decisional process. Discriminability (d') reflects the ability to perceive the difference between the distribution of perceptual evidence on "same" trials and the distribution of perceptual evidence on "different" trials. In other words, d' quantifies how good someone is at discriminating between same and different target parts. But, a same/different decision also depends on where a participant places the criterion for when to respond "same" or "different" — how much perceptual similarity is required before a "same" response is made? One purpose of SDT is to derive a measure of discriminability based on hits (e.g., correctly responding "same") and false alarms (e.g., incorrectly responding "same") that is independent of criterion (also called response bias).

As described above, because predictions cannot be made about "different" trials in the partial design, the alignment effect is based on accuracy on "same" trials, or, in SDT terminology, hits. Thus, the alignment effect refers to a higher hit rate for misaligned vs. aligned trials. However, there are two different ways that hits could be influenced by alignment: 1) if discriminability changes, or 2) if response bias changes. Therefore, changes in hits due to misalignment could arise due to changes in response bias (e.g., a tendency to respond "same" more often on misaligned than aligned trials) rather than perceptual discriminability (e.g., greater ability to discriminate between same and different target parts when face halves are misaligned). Critically, it is impossible to know which of these possibilities accounts for changes in hit rate when only "same" trials are analyzed. Of course, the partial design includes "different" trials, so one option is to complement the hits measure with a measure of false alarms (responding "same" on different trials) in the aligned and misaligned conditions (see Konar et al., 2010; Richler, Cheung & Gauthier, 2011b). This ensures that any difference in response bias driven by alignment is accounted for. However, as described next, there is another potential source of response bias in the partial design that is not addressed by this (or any other) solution.

The Problematic Response Bias in the Composite Task

The response bias we are concerned with here is one that is most salient when considering the complete design, because it has to do with congruency. The response bias associated with congruency (how willing one is to respond "same" as a function of congruency) can vary independently of the congruency effect (difference in discriminability between congruent and incongruent trials) itself. For example, Richler, Mack, Gauthier, & Palmeri (2009) manipulated the presentation time for either the study or test face in a complete design version of the composite task that included both congruent and incongruent trials, always aligned. As shown in Figure 2 (top panel), at the most rapid presentation times (<200ms), whether the presentation time of the study or test face was manipulated influenced response bias, with participants more likely to respond "same" when the test face was presented briefly, and more likely to respond "different" when the study face was presented briefly. For presentation times that were longer than 200ms (which is more consistent with the presentation times typically used in composite task experiments) response bias was related to congruency, with participants more likely to respond "different" on incongruent trials. Importantly, these changes in response bias across presentation times did not influence the congruency effect in d'. As shown in Figure 2 (bottom panel) congruency effects were observed for all presentation times regardless of whether the presentation time of the study or test face was manipulated. In other words, even though whether the study or test face was presented briefly led to differences in response biases associated with congruency, this manipulation did not influence discriminability.

This response bias associated with congruency is a confound in the partial design because the irrelevant part is always "different". Therefore, all "same" trials are incongruent and all "different" trials are congruent (see Figure 1) — correct response is fully confounded with congruency. Although congruency is not a factor of interest in the partial design, its influence on response bias (e.g., Richler, Cheung & Gauthier, 2011a,b; Richler, Mack, et al., 2009; 2011) cannot be factored out, even when d' is used; all "same" trials (hits) are incongruent, and all "different" trials (false alarms) are congruent. As we have argued elsewhere, even if congruency is of no interest to authors who use the partial design, it remains a potential source of variability in their data (Richler & Gauthier, 2013). Conversely, if researchers *are* interested in the response biases associated with holistic processing, they cannot be measured in the partial design because of this confound. Next, we demonstrate that this confound and its influence on response bias influence the results obtained with the partial design.

Does this Problematic Response Bias Lead to Different Results between Designs?

A difference in response bias between congruent and incongruent trials has been replicated across a number of studies (Richler, Gauthier et al., 2008; Richler, Mack et al., 2009; 2011). Furthermore, the difference in response bias based on congruency is often modulated by an interaction with alignment (Cheung et al., 2008; Richler, Gauthier et al., 2008; Richler, Mack et al., 2011), which itself is associated with significant response biases (Cheung et al., 2008; Richler, Tanaka et al., 2008). By themselves, these findings do not unequivocally demonstrate that there is a problem with the partial design. However, changes in response biases across experimental conditions have been quantitatively linked to differences in the

partial design measure of holistic processing across those conditions (Cheung et al., 2008; Richler, Cheung & Gauthier, 2011a,b; Richler, Mack et al., 2011). Critically, this association between response bias and the partial design leads to different results and conclusions about holistic processing than those obtained with the complete design.

For example, Goffaux & Rossion (2006) claimed that holistic processing is supported by low spatial frequencies because they observed a larger alignment effect for faces that were filtered to preserve only low- vs. high-spatial frequency information. Looking at only the partial design trials (which can be extracted from the complete design), Cheung, Richler, Palmeri & Gauthier (2008) replicated this finding. However, using the complete design data for the same participants, they found that response bias also varied across spatial frequency conditions. These differences in response bias measured in the complete design predict the observed differences in the alignment effect between high- and low-spatial frequency faces obtained in the partial design: An additional analysis of the data from Cheung et al. (2008) reveals a significant correlation between the interaction in response bias (congruency × alignment × spatial frequency) and the difference in the magnitude of the alignment effect between high- and low-spatial frequency conditions ($r_{21}= -.43$, $p = .05$). In contrast, holistic processing assessed with the complete design measure that is independent of response biases was unaffected by spatial frequency content[3]. Therefore, with the same participants and the same data set, different conclusions were drawn about holistic processing depending on which measure was used because the partial design measure of holistic processing was modulated by response biases that varied between the spatial frequency conditions.

As another example, studies using the partial design have suggested that inversion eliminates or at least dramatically reduces holistic processing (Goffaux & Rossion, 2006; 2007; Michel et al., 2006; Rossion & Boremanse, 2008; Susilo et al., 2013). However, a complete design study (Richler, Mack et al., 2011) showed that although performance was quantitatively worse for inverted vs. upright faces, the magnitude of the congruency × alignment interaction indicative of holistic processing in the complete design did not differ between orientations. These differences between the partial and complete designs of the composite task were again linked to differences in response bias between upright and inverted faces that affected the partial design measure: response bias based on congruency, alignment, and orientation was correlated with differences in the partial design measure of holistic processing between upright and inverted faces. Such a correlation was not observed for the complete design measure (Richler, Mack et al., 2011).

In summary, although congruency is not a factor of interest in the partial design, it nonetheless influences the results obtained using the partial design because of its effect on response bias. Response biases in the composite task may be of interest, but they cannot be quantified in the partial design because of the confound between congruency and correct response. In contrast, the measure of holistic processing in the complete design is

---

[3]There may be an advantage for holistic processing in LSF under some conditions: for instance, Goffaux (2009) found larger holistic processing for LSF than HSF faces in a complete composite design where there was no line or cue separating the attended from unattended region. Thus, it is possible that a difference in spatial uncertainty between spatial frequency conditions influenced the spread of attention.

independent of these response biases, and both holistic processing and response biases can be measured. Consequently, the results from the partial and complete designs of the composite task are often inconsistent with each other, even when they are based on the same participants and data set.

Partial vs. Complete Design: Testing Construct Validity

The studies reviewed above demonstrate that the design of the composite task used matters: different versions lead to qualitatively different results due to influences of response bias in the partial design. But they are not informative about which design is *better* for measuring holistic processing. For example, it could be that the biases that influence the partial design and holistic processing are one and the same (Rossion, 2013). Indeed, in the composite illusion, when the same top face half is paired with a different bottom face half (incongruent trial), the tops appear more different…so it may not be surprising that participants are more likely to respond "different" on incongruent trials.

This argument is reasonable for response biases based on stimulus properties (e.g., spatial frequency content, inversion) because there are theoretical reasons to suggest that these might influence holistic processing (see Goffaux & Rossion, 2006; Rossion, 2008). But response biases can also be influenced by top-down strategy, which should not influence holistic processing. If stimulus-driven response biases can impact measures of holistic processing in the partial design, isn't it possible that this task is also susceptible to strategy-driven response biases?

This was the question posed by Richler, Cheung, & Gauthier (2011a). In that study, participant strategy was manipulated either explicitly through instructions, or implicitly by varying the proportion of same/different trials. In both cases changes in response bias that corresponded with strategy (i.e., tendency to respond "same" more often when 75% of the trials were "same") influenced the measure of holistic processing in the partial design but not the complete design. This was particularly striking when strategy was manipulated explicitly. Participants were either told that 75% of the trials in the experiment would be "same" or that 75% of the trials would be "different", when in fact the proportion of same/different trials was always 50%. Nonetheless, response biases created by these instructions modulated whether evidence for holistic processing was observed in the partial design, with no evidence for holistic processing when participants were told that there would be more "different" trials in the experiment. In other words, participants' beliefs alone influenced whether or not faces were processed holistically according to the partial design measure! Such a finding is inconsistent with all theories of holistic processing because they all posit that holistic processing is obligatory —participants are explicitly instructed to selectively attend, yet they are unable to do so, either because faces are encoded into a holistic representation (Tanaka & Farah, 1993) or because expertise has increased the automaticity of this perceptual strategy (Richler, Wong, & Gauthier, 2011). Thus, the evidence suggests that the partial measure of holistic processing should be abandoned because it is tapping into something strategic, rather than a visual phenomenon.

Interim Conclusion

We cannot continue to use both of these inconsistent measures of holistic processing as if they were measures of the same thing. Results from the two designs will not converge, and they cannot be combined in formal or informal meta-analytic reviews because it is impossible to determine whether differences between studies are due to purposeful experimental manipulations or are an artifact of the version of the composite task being used. More critically, any result obtained with the partial design must be questioned because the effects could be driven by response biases and not holistic processing itself. We contend that the complete design measure of holistic processing should be preferred because the partial design measure 1) confounds manipulations of interest with complex response biases; and 2) has poor construct validity, as it can be modulated by participant beliefs.

Meta-Analysis

In the next section, we present a meta-analysis of holistic processing of faces in the complete and partial designs of the composite task. Our first goal was to compare the effect sizes obtained from the complete design (congruency × alignment interaction in d') and partial design (difference in percent correct on "same" aligned vs. misaligned trials). Our second goal was to analyze the size of the correlation between the congruency effect on aligned and misaligned trials in the complete design. This is relevant for measurement decisions in individual differences studies and for accurately calculating power in studies using this task, as such calculations take the correlation between conditions into account.

Search Methods

Studies for the meta-analysis were collected by the first author. A literature search in Google Scholar was conducted with the terms "holistic processing" and "composite task". Methods sections were examined to determine study eligibility. Datasets were included if the complete composite design was used, both congruency and alignment were manipulated in the experiment, the dependent variable was sensitivity, and participants were normal adults. Note that although we only had one rater for study inclusion, these are very objective, dichotomous criteria. Included papers were published or online-first by December 31, 2012[4]. Unpublished data were collected prior to this date. If the data necessary for calculating effect size (η^2_p) were not available in the published paper or if the published data were in RTs, corresponding authors were contacted and asked to provide the full ANOVA table for their analysis in sensitivity (the response rate to such queries was 100%). We also sought unpublished and submitted datasets by contacting colleagues in the field who study holistic processing of faces using the complete design. Unpublished and published data were treated in the same manner. Study quality was not assessed, nor were any moderator/ mediator variables analyzed.

To analyze the effect size of holistic processing in the partial design, we calculated the partial design measure of holistic processing (difference in hit rate on same-incongruent misaligned vs. aligned trials) from data collected in the context of complete design studies,

---

[4]One exception is Chua, Richler, & Gauthier (in press). These data were initially included as unpublished data, but between the first submission of this manuscript and submission of the revision they were accepted for publication.

where these data were available to us. This ensures that all methodological details are the same for the partial and complete designs (e.g., instructions, subjects, stimuli etc.). Previous work has shown that the same qualitative effects are obtained when partial design trials are randomized within the context of the complete design and when they are run separately (see Richler, Cheung, & Gauthier, 2011b, supplemental materials). Most importantly, any other comparison of the partial design to different complete design studies would be complicated by many confounds, whereas this is a comparison that keeps stimuli, task parameters, and subjects constant. Moreover, this approach has the benefit of producing an estimate of the partial design effect that is not susceptible to publication bias, as the datasets being used were not selected because they gave a publishable partial design effect.

The analysis of the correlation between congruency effect-aligned and congruency effect-misaligned was calculated in a subset of studies where the raw data were available to us.

Statistical Methods

Effect sizes for holistic processing in the complete and partial design were calculated by the formula:

$$\eta^2_{\mathrm{p}} = SS_{\mathrm{effect}} / (SS_{\mathrm{effect}} + SS_{\mathrm{error}})$$

Confidence intervals were calculated by taking the square root of η^2_{p} to get an r value (Rosenthal, Rosnow, & Rubin, 2000), using an online calculator (http://vassarstats.net/rho.html) to calculate end points, and then transforming these values back to the η^2_{p} scale.

Pearson's was used as the effect size for the correlation between the congruency effect-aligned and congruency effect-misaligned. Confidence intervals were calculated using the same online calculator. We divided the datasets into tertiles based on a rank ordering of the correlations, and screened datasets in the top and bottom tertiles for bivariate outliers using Cook's D. This was judged as important to produce trustworthy estimates because the correlation between congruency effect-aligned and congruency effect-misaligned was not screened in any of these studies (as this correlation was never considered directly), and correlations are particularly sensitive to outliers. As a result, data from 20 subjects (from 17 datasets, at most two in a given sample) were excluded from this analysis.

For all meta-analyses, a mean weighted effect size was calculated using a random-effects model following the procedure outlined by Ellis (2011, Appendix 2). For these calculations, η^2_{p} were transformed to r (by taking the square root) and then to z-scores using the FISHER function in Excel.

First, heterogeneity of effect sizes was assessed by Cochran's Q (Cochran, 1954):

$$Q = \Sigma w^* (z_{\mathrm{dataset}} - z_{\mathrm{meta}})^2$$

Weights (w) were the inverse of within-study variance, which was estimated by the formula $1/(n-3)$. Thus, weights were equal to $n-3$ in this first step.

Next, we estimated between-study variance using the following formula:

$$\tau^2 = \left[Q - (k-1) \right] / c$$

Where k is the number of studies, and c is calculated by:

$$c = \Sigma w - \left(\Sigma w^2 / \Sigma w \right)$$

The new weights ($w*$) for the random-effects calculation of mean weighted effect size were equal to the inverse of within-study variance plus between-study variance:

$$w* = 1 / \left(1 / (n-3) + \tau^2 \right)$$

Confidence intervals for mean-weighted effect sizes were calculated using the standard error of the random-effects mean (square root of $1 / \Sigma w*$).

$$CI_{\mathrm{lower}} = \mathrm{mean-weighted~~effect~~size} - (1.96 \times SE)$$

$$CI_{\mathrm{upper}} = \mathrm{mean-weighted~~effect~~size} + (1.96 \times SE)$$

Finally, the mean weighted effect size and CI were transformed back to r using the FISHERINV function in Excel, and then back to η^2_p by r^2.

Evidence of publication bias was assessed by calculating the correlation (Pearson's r) between N and effect size (Schimmack, 2012).

Results

Our search resulted in 48 independent data sets (N = 1545), 16 of which are unpublished (33.3%)[5]. An additional 6 studies (12.5%) are included from manuscripts that are under review. Table 1 shows the complete list of studies included in the meta-analysis with a brief summary of some key experimental conditions.

Effect sizes (η^2_p) for holistic processing in the complete design (congruency × alignment interaction in d') for each study are plotted in Figure 3, with marker size indicating sample size. Q_{obtained} (88.55) > $Q_{\mathrm{crit, df = 47}}$ (64.00), indicating homogeneity of effect sizes (see Higgins, 2008, and Coory, 2010, for a discussion of heterogeneity in meta-analysis). The mean weighted effect size across studies was $\eta^2_p = .32$ (95% CI: .26, .38). The correlation between N and effect size was not significant ($r_{48} = -.14$, $p = .36$), providing no reason to suspect publication bias.

---

[5]Nine datasets (19%) are from other labs. This accounts for 34.6% of the published studies included in the meta-analysis. Naturally, we had greater access to our own unpublished and submitted data, but we also have been the primary advocates of using the complete design, so these numbers likely reflect the reality of the existing data with this task.

The partial design meta-analysis included 28 independent datasets (marked with asterisks in Table 1; N = 972), 12 of which are unpublished (42.9%). An additional 6 studies (21.4%) are included in manuscripts that are under review.

Effect sizes for holistic processing in the partial design for each study (η^2_p) are plotted in Figure 4, with marker size indicating the size of the sample. $Q_{obtained}$ = 166. 98 ($Q_{crit, df = 27}$ = 40.11) when we included all the datasets marked with an asterisks in Table 1. However, inspection of Figure 4 suggests one clear outlier (marked '×'). Removing this data point leads to a much smaller Q statistic ($Q_{obtained}$ = 80.12, $Q_{crit, df = 26}$ = 38.89) suggesting that this study was introducing considerably more variability than the other studies. We therefore excluded this study from the meta-analysis of the partial design effect. The mean weighted effect size for the alignment effect was η^2_p = .14 (95% CI: .07, .22). The correlation between effect size and N was not significant ($r_{27} = -.31$, $p = .12$), which was expected given these studies were written up with a focus on a different effect.

For better comparison with the complete design, we also calculated the mean weighted effect size for holistic processing in the complete design for the subset of studies included in the partial design meta-analysis. The mean weighted effect size was remarkably similar to that obtained from the full set of studies (η^2_p = .35, 95% CI: .28, .41). Again the correlation between effect size and N was not significant ($r_{27} = -.22$, $p = .27$). There was no significant correlation between the partial and complete design effect sizes ($r_{27} = .27$, $p = .18$), consistent with the prior discussion that they measure different constructs. Since response bias across subjects has been found to predict the magnitude of the partial design holistic processing measure (Richler, Cheung, & Gauthier, 2011a,b; Richler, Mack et al., 2011), we asked whether a similar association would exist across studies. Indeed, response bias (congruency × alignment interaction in criterion) predicts the effect size of holistic processing in the partial design ($r_{27} = -.38$, $p = .048$). This correlation is not significant for the complete design ($r_{27} = -.19$, $p = .34$).

Finally, the meta-analysis of the correlation between the congruency effect-aligned and congruency effect-misaligned included 39 independent datasets (marked with carets in Table 1; N = 1223), 16 of which are unpublished (41.0%). An additional 6 studies (15.4%) are included in manuscripts that are under review.

Effect sizes for this correlation (*r*) for each study are plotted in Figure 5, with marker size indicating the size of the sample. The mean weighted effect size was *r* = .24 (95% CI: .16, .32)[6]. $Q_{obtained}$ (74.91) > $Q_{crit, df = 38}$ (53.38) indicating heterogeneity of effect sizes. The correlation between effect size and N was not significant ($r_{39} = -.23$, $p = .16$).

Discussion

Our meta-analysis revealed variability in the effect size of holistic processing in the complete design across experiments. This variability is not surprising, since our meta-analysis has to include sampling variability present in any field of research. Sampling variability accounts for the fact that sometimes studies that use the same methods do not

___

[6]The mean weighted effect size with outliers included is *r* = .29 (95%CI = .20, .38).

produce the same qualitative results, especially when alpha levels are treated like a cliff and effect sizes are ignored. Such variability cannot be eliminated, and it is not as problematic as differences in results due to the incommensurability of measurements that we argue have complicated the literature on the composite task.

We also found that the holistic processing effect measured in the complete design is moderate in size and robust. A power analysis (G*Power 3.1) using the meta-analysis correlation between congruency effect-aligned and congruency effect–misaligned indicates that the congruency × alignment interaction requires 13 subjects for 95% power (alpha = . 05). Notably, the complete design effect is almost three times larger than the partial design effect. It has been argued that the partial design measure is better suited for children and clinical populations because it requires only half of the trials (Rossion, 2013; but see Richler & Gauthier, 2013, appendix). However, these results suggest that due to the small effect size, more subjects need to be included in partial design studies to achieve adequate power. Thus, for populations in which subject recruitment is more challenging, these data suggest another reason that the complete design should be preferred.

Of course, there may be aspects of studies designed to only measure the partial design that may increase the partial design effect size (see Rossion, 2013). However, it remains an open question as to whether these same factors also increase the magnitude of the complete design effect. For now, this is the only fair meta-analytic comparison of the partial and complete designs that we can think of, since here the data are from the same experiments, keeping all design factors constant across the partial and complete design calculations. The results suggested no publication bias in the complete design and indeed we included both our published and unpublished data. Despite the fact that these measures were calculated from the same datasets, there was no substantial correlation between the partial and complete design effect sizes, providing further evidence that they are not in fact measuring the same thing. In contrast, and as might be expected based on earlier sections of our review, response bias predicts the effect size of holistic processing in the partial design, but not the complete design.

In the third part of our meta-analysis, we calculated the effect size of the correlation between congruency effect-aligned and congruency effect-misaligned. This is required for accurately calculating power for the complete design when holistic processing is measured as the subtraction of the misaligned from the aligned conditions (see above). More generally, these results are important for decisions on how to best measure holistic processing, and what conditions are needed in different experimental situations. The absent or much reduced congruency effect in the misaligned compared to aligned condition, on average, has been theoretically important to argue that holistic processing is specific to the aligned configuration, which is why the congruency × alignment interaction has been the preferred measure in studies where data is analyzed at the group-level. However, the correlation between the two congruency effects informs how holistic processing should be measured in an individual differences framework.

Measurements with the composite task are often reported to have low reliability, and this is the case regardless of whether performance on misaligned trials is subtracted from

performance on aligned trials, as is the case in most published work, or whether it is regressed out, as advocated more recently (e.g., DeGutis et al., 2013; Ross et al., under review). Such low reliability may not be a problem in group studies (Thomas & Zumbo, 2012), but it limits the precision of estimates of holistic processing in individual differences work. The results of our meta-analysis suggest that to the extent that one is measuring holistic processing under conditions known to be specific to the aligned condition, collecting misaligned data may not be necessary – in fact, a much better use of participants' time may be to collect more data in the aligned condition to improve power. Indeed, we find that the shared variance between the congruency effect on aligned and misaligned trials is very small (6%). This is consistent with the idea that misaligned face parts are not processed holistically. Although a baseline such as the misaligned condition was important in the development of this measure and may remain critical when new conditions or categories are tested, we propose that the misaligned condition can be dropped in individual differences measures of holistic processing. If the congruency effect in the misaligned condition does not measure something that is a substantial confound in estimates of holistic processing in the aligned condition (which would motivate a regression approach), there is no need to measure it. Regressing variance in the misaligned condition out of that in the aligned condition does little more than regressing out anything else (verbal IQ, shoe size…) that shares little variance with holistic processing. Thus, our meta-analysis results suggest that when the goal is to quantify holistic processing in situations where it is well known that misaligned stimuli do not yield a congruency effect (e.g., normal populations with standard face stimuli), the best approach would be to increase the number of aligned trials and not collect data from misaligned trials[7].

Less is More with Holistic Processing

In the final part of our paper, we provide the first review of holistic processing based only on the results obtained from the complete design. The current review focuses on holistic processing of identity, but the complete design of the composite task has also been used to investigate holistic processing in other types of face judgments, such as emotion (Tanaka et al., 2012) and social judgments (Todorov et al., 2010).

Studies that have used the partial design represent a sizeable amount of literature, which comprises a valuable contribution of theoretical ideas and interesting experimental manipulations. However, our review and meta-analysis provide clear convergent evidence that the two versions of the composite task do not measure the same construct. Because the partial design proves to be sensitive to response biases at the level of individual subjects (Richler, Cheung, & Gauthier, 2011,a,b; Richler, Mack et al., 2011) and across different studies, we do not believe that including partial design results in our theoretical review is appropriate. Insofar as the sole empirical basis for a claim is the partial design, the evidence

---

[7]As an example, additional analysis of the data from Richler, Cheung & Gauthier (2011b) reveals that in a multiple regression analysis with all predictors entered simultaneously the congruency effect on aligned trials is a significant predictor of face recognition performance (B = 5.63, SE = 2.44, t = 2.31, p = .027) while the congruency effect on misaligned trials is not (B = −5.72, SE = 3.11, t = −1.84, p = .074). This pattern of results is replicated in the dataset from Experiment 3 of McGugin, Richler et al. (2012; congruency effect aligned: B = 6.65, SE = 3.18, t = 2.09, p = .04; congruency effect misaligned: B = −3.01, SE = 3.88, t = −.78, p = .44). Thus, the misaligned conditions do not contribute to our understanding of how individual differences in holistic processing are related to face recognition abilities.

presented in the previous sections suggests that this claim needs to be empirically re-evaluated, because the effects of holistic processing and response biases cannot be disentangled.

The complete design's validity has survived the same tests we imposed on the partial design. It has face validity, measuring the extent to which one is able to ignore a task-irrelevant part of an object when it is in a familiar configuration. Its construct validity is also good, as it captures the difference between faces and objects that has been of interest for decades: the complete design of the composite task consistently leads to positive evidence of holistic processing for upright faces (e.g., Cheung et al., 2008; Farah et al., 1998; Richler, Mack et al., 2009; 2011; Richler, Tanaka et al., 2008) and consistently shows little evidence of holistic processing for non-face objects in novices (Farah et al., 1998; Richler, Mack et al., 2011). Adding to its discriminant validity, it is not affected by manipulations or normal variability in response bias, which is not theoretically expected to influence holistic processing. The next question, then, is what else do we know about holistic processing that can help us understand this difference between faces and objects? A summary of the conclusions from our review of what we know about holistic processing from the complete design only, with some outstanding questions, is presented in Table 2.

Holistic Processing & Expertise

We start with the debate that motivated this review, and this line of research more generally, in the first place: are faces "special", or is performance with faces the result of experience and expertise? Critically, although holistic processing of objects is not observed in novices, it is observed for experts: Holistic processing has been observed for cars in car experts (Bukach et al., 2010; Gauthier et al., 2003), chess boards in chess experts (Boggan et al., 2012), short sequences of music notes in music reading experts (Wong & Gauthier, 2010), and for English words in English readers (Wong et al., 2011). Thus, holistic processing may be observed for faces in the general population not because holistic processing is a special mechanism for face processing, but rather because, due to a lifetime of experience recognizing and identifying faces, most people are "face experts". Indeed, holistic processing in real-world experts is specific to the category subclass with which expertise was acquired. For example, experts with modern cars show holistic processing for modern cars, but not antique cars (Bukach et al., 2010), and native English readers show greater holistic processing for real words than pseudo-words (Wong et al., 2011).

One exception to this general pattern of results is a study by Hsiao & Cottrell (2009). Here, novices, but not experts, showed evidence for holistic processing with Chinese characters. However, when a different response deadline was used for novices vs. experts to get expert performance off ceiling, Wong, Bukach et al. (2012) found robust holistic processing for Chinese characters in experts and evidence that failures of selective attention in the two groups were driven by different mechanisms.

Training studies with novel objects are particularly informative about what aspects of face perception promote the development of holistic processing because specific hypotheses can be tested and the amount or type of experience can be manipulated. Indeed, amount and quality of experience appear to mediate holistic processing in real-world experts. For

example, greater holistic processing is observed for native English readers vs. Chinese readers who learned English as a second language and therefore have less experience with English words (Wong et al., 2011). Training studies have shown that subordinate-level categorization is the critical component leading to holistic processing. Experience individuating novel objects leads to holistic processing (Gauthier & Tarr, 2002; Gauthier et al., 1998; Wong et al., 2009a), while other types of training with equivalent exposure do not (Wong et al., 2009a).

Finally, holistic processing in experts correlates with face-specific neural markers, such as activity in the right FFA (Gauthier & Tarr, 2002; Wong et al., 2009b) and the N170 ERP component (Gauthier et al., 2003). It should be noted that although holistic processing has been observed with words (Wong et al., 2011), words do not activate the right FFA, but rather the homologous *left* fusiform gyrus (visual word form area, see McCandliss, Cohen & Dehaene, 2003 for a review). In fact, literacy (e.g., reading experience) leads to stronger lateralization of letter strings to the left hemisphere and faces to the right hemisphere (Dehaene et al., 2010), suggesting that hemispheric organization of face and word recognition do not develop independently (Dundas, Plaut & Behrmann, 2013). Given that reading is associated with different task demands than face-like expertise (e.g., McCandliss et al., 2003; Wong et al., 2009a), the nature of holistic effects may differ between categories for which individuation is required (and which recruit the FFA) and those that involve a different kind of perceptual expertise (Wong & Gauthier, 2007). Alternatively, holistic processing may reflect a more general mechanism that emerges with expertise, and the correlation between holistic processing and brain activity may be mediated by a third variable (i.e., amount of expertise predicts both the magnitude of holistic effects and neural activity in relevant brain areas). More work is required to determine whether holistic processing of words is qualitatively similar to holistic processing of faces and other domains of individuation expertise like cars or birds.

In any case, the correlation between holistic processing and FFA activity for objects as a function of expertise is theoretically consistent with findings for other hallmarks of face perception besides holistic processing that also depend on expertise. These include a larger visual short-term memory capacity (Curby & Gauthier, 2007; Curby, Glazek & Gauthier, 2009), a greater sensitivity to manipulations of spatial frequency content (McGugin & Gauthier, 2009) and mandatory processing when these objects are presented as distractors under perceptual load (Ro, Friggel & Lavie, 2009). In sum, face-specific effects are typically obtained with non-face objects of expertise, and objects of expertise can recruit the FFA (Bilalic et al., 2011; Gauthier et al., 1999; 2000, 2003; Xu, 2005; McGugin, Gatenby et al., 2012; Wong, Folstein & Gauthier, 2012). However, not all visual learning recruits the FFA or even the fusiform gyrus (Brants, Wagemans, & Op de Beeck, 2011; Op de Beeck & Baker, 2010), and given the available evidence, one conjecture is that it is expertise with individuation of objects from a homogenous category, in particular when this expertise increases holistic processing, that recruits the fusiform gyrus (Wong et al., 2009a,b).

Mechanisms of Holistic Processing

The evidence reviewed above suggests that individuation experience gives rise to holistic processing, but what are the mechanism(s) that explain this effect? Surprisingly, although the holistic encoding hypothesis is the most common description of the mechanism underlying holistic processing, it has rarely been tested empirically, and is often supported by the conjecture that holistic processing is "fast" or "early" and must therefore have a perceptual basis in encoding (e.g., Jacques & Rossion, 2009; Rossion, 2013). Behaviorally, evidence in favor of holistic encoding has only been directly provided by a study where inverting faces presented during a learning phase eliminated holistic processing measured at recognition, where subjects judged whether or not the top halves of upright composite faces were part of a previously studied face (Boutet et al., 2002). However, the interpretation of these results is questionable. First, inversion disrupts configural processing (e.g., Leder & Bruce, 2000; Searcy & Bartlett, 1996), while it only seems to delay holistic processing (Richler, Mack et al., 2011). Second, in the composite task manipulations of alignment only disrupt holistic processing when applied to the test face, not the study face (Richler, Tanaka et al., 2008; see Figure 6; see also Cheung & Gauthier, 2010, for evidence that manipulations at test are more disruptive to holistic processing than manipulations at study), which is inconsistent with the idea that holistic processing results from a comparison to a holistic template of the study face.

Studies that include an inversion manipulation are also incompatible with a template account of holistic processing. For example, performance at matching parts of inverted faces is above chance when presentation times are very rapid (50ms and 183ms), but interference indicative of holistic processing is not observed until presentation times are longer (e.g., 800ms presentation, Richler, Mack et al., 2011). In contrast, holistic processing of upright faces is observed earlier (183ms presentations). Above chance performance only requires encoding the target part, whereas interference requires that the irrelevant part be encoded as well. Thus, one explanation for this pattern of results is that the target parts of inverted faces were successfully encoded at the more rapid exposure durations, but less efficient processing due to inversion delayed encoding of the non-attended face part, resulting in no interference. Presumably upright faces are encoded in the same manner as inverted faces, but the efficiency with which upright faces are processed does not allow the resolution necessary to observe this effect; although they may be encoded separately, increased processing efficiency for upright faces allows both the target and distractor parts to be encoded rapidly, leading to interference even when presentation times are very brief. Thus, in contrast to a template account, face parts appear to be encoded and represented independently, but not treated independently during perceptual decisions.

Other work supports this idea that upright and inverted faces rely on the same qualitative processing strategy, but differ in terms of processing efficiency (Loftus & Dillon, 2004; Riesenhuber et al., 2004; Sekuler et al., 2004; Valentine & Bruce, 1988; Willenbockel et al., 2010). Indeed, the FFA responds preferentially to both upright and inverted faces (Kanwisher, Tong & Nakayama, 1998) compared to other objects, although activity is reduced for inverted vs. upright faces (Gauthier et al., 1999; Yovel & Kanwisher, 2005). The fact that holistic processing takes longer to emerge for inverted vs. upright faces is also

consistent with the delay in the N170 ERP component for inverted vs. upright faces (although the delay is only on the order of 10 ms on that component; e.g., Bentin et al., 1996; Eimer, 2000; Sagiv & Bentin, 2001; Rossion et al., 2000) and with other behavioral work showing an encoding advantage for upright vs. inverted faces and cars in car experts vs. novices (Curby & Gauthier, 2009; see also Curby & Gauthier, 2007; Curby et al., 2009).

An alternative to the holistic encoding hypothesis is that holistic processing is the outcome of an attentional strategy that has become automatized with experience (Richler, Palmeri & Gauthier, 2012; Richler, Wong & Gauthier, 2011). This theory is supported by evidence that contextual manipulations that induce bottom-up attentional biases penetrate holistic processing, even though holistic processing itself does not necessitate general attentional resources (Boutet et al., 2002). For example, Gao and colleagues (2011) induced global or local processing biases by having participants perform either a local or global task (blocked) with Navon letters (Navon, 1977) prior to each trial of a composite face task. Holistic processing of faces was larger following a task that required attention to the global properties of the Navon letters. Similarly, inducing a negative mood, which is thought to promote a local processing bias (Basso et al., 1996; Derryberry & Reed, 1998; but see Huntsinger, 2012), leads to reduced holistic processing relative to a positive or neutral mood (Curby, Johnson & Tyson, 2012).

Interestingly, like a global Navon task, aligned faces themselves can create an attentional bias that influences subsequent processing of non-face objects. In a task where a face composite task and a novel object composite task were interleaved, congruency effects were observed for novel objects if they were preceded by an aligned face — a stimulus that is processed holistically — but not a misaligned face (Richler, Bukach & Gauthier, 2009). This finding is difficult to explain by invoking a face template: how would the use of a particular template for aligned faces impact processing of a subsequent object that does not share the same configuration of features? Instead, these results suggest that holistic processing in expertise is a consequence of an automatized perceptual strategy; the holistic processing strategy that was automatically recruited for the aligned face stimulus could not be "turned off" in time to process the novel object, resulting in an inability to successfully selectively attend to one part of the object.

The idea that holistic processing of faces can be understood within the context of domain-general attentional processes is supported by two recent studies. In Curby, Goldstein & Blacker (2013), face parts were always presented in an aligned format. Critically, square regions surrounding the two face halves were either the same color and aligned, or different colors and misaligned. Remarkably, this manipulation led to a decrease in holistic processing that was similar in magnitude to that observed when face parts themselves are misaligned. In other words, discouraging the grouping of face parts by disrupting classic Gestalt cues of similarity (the color of the background) and continuity (alignment of color blocks) reduced holistic processing in the same manner as physically misaligning the face parts. Even more compelling, when faces were inverted the results were strikingly similar to those of Richler, Mack et al. (2011): a congruency × alignment interaction was only observed for upright faces and not inverted faces at rapid exposure durations (200ms). But,

when presentation times were longer, overall sensitivity was lower for inverted faces, but the pattern was the same as that observed for the upright faces.

In another study, Chua, Richler, & Gauthier (in press) trained subjects to individuate faces from two novel races, Lunaris and Taiyos. During the individuation training, diagnostic information was presented in the top face half for one race (e.g., Lunaris), and the bottom face half for the other race (e.g., Taiyos). After training, participants completed a composite task with Lunari-Taiyo composites that were either made with diagnostic (e.g., Lunari top/ Taiyo bottom) or non-diagnostic (e.g., Taiyo top/Lunari bottom) parts. Critically, only composites made from diagnostic parts were processed holistically, even though these Lunari/Taiyo parts had never been experienced together. These results, in which holistic effects were obtained after participants learned to attend to diagnostic parts, raise the intriguing possibility that faces are processed holistically because of part-based attention to several parts, which is strengthened and automatized as a result of experience. This represents an exciting avenue for future work.

Holistic Processing & Competition

In contrast to studies where holistic processing of faces can transfer to other subsequently presented objects that are not themselves processed holistically (Gao et al., 2011; Richler, Bukach & Gauthier, 2009), other studies find that holistic processing of one stimulus can inhibit or reduce holistic processing of subsequently presented objects when those objects also exhibit automatic holistic processing. For example, in car experts (who process cars holistically), holistic processing of faces was reduced when face and car composite tasks were interleaved (Gauthier et al., 2003); face-selective neural responses and behavioral performance were reduced in a prosopagnosic patient as a function of training to individuate non-face objects (Behrmann et al., 2005). These results have been interpreted to suggest that holistic processing of faces and objects of expertise recruit a common resource, with other work suggesting that the locus of this competition is perceptual (Cheung & Gauthier, 2010).

Holistic Processing vs. Strategic Congruency Effects

Of course, failures of selective attention are not unique to face perception or expertise. Indeed, Stroop and flanker tasks bear many similarities to the composite task: task-irrelevant information interferes with performance despite instructions to selectively attend to a target stimulus or stimulus property. Thus, one challenge is to distinguish between failures of selective attention in the composite task that arise because of holistic processing in experts and those that could arise for other reasons, even in novices.

Wong & Gauthier (2010) examined holistic processing of musical notation in experts and novices by having participants judge whether a cued target note was the same or different in two sequentially presented note sequences. Critically, target position (central or peripheral) and target distribution (mostly in the center, mostly in the periphery or evenly distributed) were manipulated, such that focusing on certain note positions would be an advantageous strategy. Failures of selective attention were observed in both experts and novices. However, unlike the experts, failures of selective attention in novices were modulated by the distribution of target position and were more pronounced when the target appeared in an

unexpected location. This suggests a role of strategy: if novice participants were devoting more attention to the location where the target appeared more frequently, then when the target appeared in the unexpected location the distractor would appear in the relatively more attended location, leading to interference. This strategic modulation of failures of selective attention in novices was observed both when target distribution was manipulated implicitly (i.e., subjects were not told that targets were more likely to appear in the central or peripheral positions) and explicitly (i.e., subjects were told that targets were more likely to appear in the central positions). Thus, unlike experts for whom holistic processing is automatic and relatively impenetrable to manipulations of strategy (Richler, Cheung & Gauthier, 2011a), failures of selective attention in novices can be induced by task contexts that explicitly or implicitly affect strategy.

More subtle task parameters can also influence strategic failures of selective attention in novices. For faces, congruency effects are larger for aligned vs. misaligned test faces, and this interaction is not influenced by whether the study face (i.e., first object in the sequential matching sequence) is aligned or misaligned, nor by whether study aligned and study misaligned trials are blocked or randomized (Richler, Tanaka et al., 2008). In contrast, study alignment influences failures of selective attention in novices. Specifically, congruency effects are observed for novel objects in novices if the study object is misaligned, and this interference spreads to all conditions if study-misaligned and study-aligned trials are randomized (Richler, Bukach & Gauthier, 2009). These results are consistent with the strategic use of a large attentional window throughout the experiment; for novices, when study items are misaligned, a larger attentional window is required to encode both parts of the study object (necessary in this study because participants did not know whether the target part would be the top or bottom until the test object was presented), but this strategy puts the task-irrelevant part of the test object within the scope of attention as well. For experts, this strategy is not necessary for successful performance because expertise increases the efficiency of encoding (Curby & Gauthier, 2009).

In summary, although holistic processing in experts can be modulated by manipulating bottom-up attentional resources (e.g., Gao et al., 2011; Curby et al., 2013), it is not susceptible to contextual modulations that influence strategy and failures of selective attention in novices (for a review see Richler, Wong & Gauthier, 2011).

Holistic Processing and Face Recognition Deficits

Deficits in face processing have been reported in several clinical populations. However, in most cases it is unclear what mechanism or process is impaired. This is even the case for prosopagnosia, an impairment defined by an inability to recognize faces! Only a single study to date has examined holistic processing in prosopagnosia using the complete design. Bukach and colleagues (2006) found that patient LR, with acquired prosopagnosia, exhibited holistic processing of faces that was comparable to controls. However, further examination of performance on a task that measured configural processing suggested that the patient had difficultly encoding multiple face parts, particularly under time constraints. These results suggest that prosopagnosia can influence face recognition performance through an impairment encoding multiple face features; ironically, even though prosopagnosics may

exhibit obligatory attention to the whole face (holistic processing), they may be unable to successfully extract relevant details from the whole face. This is consistent with the idea that face expertise leads to an encoding advantage, where more information, even more objects, can be encoded in a brief amount of time (Curby & Gauthier, 2007; 2009). An important caveat is that acquired prosopagnosia is a very heterogeneous disorder and the results obtained with LR may not generalize to other patients.

Perhaps somewhat more homogeneous as a population are individuals with developmental (also called congenital) prosopagnosia, but they have only been tested in partial design studies (e.g., Avidan et al., 2011; Palermo et al., 2011). Thus, future work is necessary to examine whether face recognition deficits in developmental prosopagnosia are due to the same type of impairment as acquired prosopagnosia, or whether holistic processing is impaired in these individuals. It may be that developmental prosopagnosia results from a failure to acquire expert processing strategies such as holistic processing. In another experiment, Chua et al. (in press) showed that subjects who were only trained with one novel race where diagnostic information was in one face half did not show evidence for holistic processing. Thus, a more general deficit in global attention in developmental prosopagnosia may result in attention to single face parts, leading to no holistic processing, similar to subjects trained on only one diagnostic part in Chua et al. In contrast, in acquired prosopagnosia these already-acquired processes may be rendered less effective.

Finally, there is evidence that face recognition deficits in Autism Spectrum Disorder are related to impaired holistic processing, consistent with the proposal that face recognition deficits in autism arise due to a failure to acquire perceptual expertise with faces because of reduced attention to faces (Schultz, 2005). Gauthier, Klaiman & Schultz (2009) tested adolescents with autism and found congruency effects that were not modulated by test face alignment. Critically, in this experiment study faces were also aligned or misaligned, a design that has been found to produce contextual effects with objects in novices (Richler, Bukach & Gauthier, 2009; see section titled "Holistic Processing vs. Strategic Congruency Effects"). Therefore, although failures of selective attention to face parts were observed in individuals with autism, face perception in this population may be more similar to the way typical participants approach a demanding part-matching task with novel objects. Future work should test individuals with autism in a composite task where study items are always aligned, which would eliminate the possibility of observing strategically-induced congruency effects.

Other-Race Effects

Like inversion, manipulations of race have well-known consequences for face recognition. The other-race effect refers to the fact that individuals are better at recognizing same-race vs. other-race faces (for a review see Meissner & Brigham, 2001). Like the comparison between upright vs. inverted faces, same- and other-race faces are both processed holistically in the few studies using the complete design (Bukach et al., 2012; Harrison, Gauthier, Hayward & Richler, under review), but there is a discrimination advantage for same-race faces (Harrison et al., under review). A similar discrimination advantage is also observed for same-race faces categorized as belonging to a social in-group vs. out-group

(Harrison et al., under review). These results suggest that, like inverted faces, overall performance decrements for other-race faces reflect superior processing efficiency for same-race faces, likely as a consequence of more experience, despite reliance on the same underlying processing mechanism. Indeed, the magnitude of the other-race effect in holistic processing (differences in holistic processing between same- and other-race faces) is correlated with individual differences in the amount of individuation experience with the other-race, such that the other-race effect is smaller for participants with more other-race individuation experience (Bukach et al., 2012). Therefore, more experience with other-race faces may improve processing efficiency that ironically hinders performance when selective attention to face parts is needed.

Conclusion

The argument developed in the first part of this review is, understandably, likely to meet resistance in the field of face processing. Many times when researchers debate the best measure of a construct, they do so on the basis of quantitative differences in reliability or sensitivity. Here, instead, we argue that one of the popular measures of holistic processing lacks validity and should no longer be used. While some still defend the partial design (Rossion, 2013), so far every published empirical comparison of the two designs has come out in favor of the complete design (Cheung et al., 2008; Richler, Cheung & Gauthier, 2011a,b; Richler, Mack et al., 2011). Indeed, an advantage of the complete design is that the partial design analyses can also be performed. One argument in favor of the partial design is that it has been used in a lot of studies and has helped the field consider a variety of interesting questions about holistic processing. But whether the conclusions of these studies are valid is impossible to assess in a vacuum. The validity of a measure needs to be assessed against a criterion, such as whether it predicts another construct that is theoretically related or whether it is robust to influences that should theoretically not influence it.

Lest we build our theories on shaky empirical foundations, prior claims based solely on partial design evidence need to be re-evaluated empirically. To be clear, it is possible, even likely, that sometimes the partial design provides an accurate estimate of holistic processing, especially in studies of aggregate effects that do not measure individual or group differences, but the fundamental problem is that it is impossible to know when response bias has contaminated the results beyond usefulness. While adopting our recommendations would require setting aside a good deal of published results, the outcome would only be a minor setback for the study of holistic processing, which could progress much faster towards an empirical body of research rooted in a more coherent construct.

Luckily, as reviewed here, there is already a sufficient body of evidence based on the complete design measure to support the notion that holistic processing is important in understanding how face processing may differ from object processing, how perceptual expertise develops, and how perceptual strategies relate to specialization in the brain. This measure is not perfect, and given that it often produces measurements that have low reliability (Ross et al., under review), a great deal of work on the measurement of holistic processing is needed to facilitate the study of individual differences in holistic processing and its role in accounting for individual differences in face recognition. We hope that this

work will encourage more authors to use the complete design to measure holistic processing to explore many important questions at the center of our understanding of face and expert processing. We also hope that some may use this analysis to extract lessons that may facilitate the creation and evaluation of even better measures for this construct.

Acknowledgments

References

Avidan G, Tanzer M, Behrmann M. Impaired holistic processing in congenital prosopagnosia. Neuropsychologia. 2011; 49:2541–2552. doi:10.1016/j.neuropsychologia.2011.05.002. [PubMed: 21601583]

Basso MR, Schefft BK, Ris MD, Dember WN. Mood and global-local visual processing. Journal of the International Neuropsychological Society. 1996; 2:249–255. http://dx.doi.org/10.1017/S1355617700001193. [PubMed: 9375191]

Behrmann M, Marotta J, Gauthier I, Tarr MJ, McKeeff TJ. Behavioral change and its neural correlates in visual agnosia after expertise training. Journal of Cognitive Neuroscience. 2005; 17(4):554–568. doi:10.1162/0898929053467613. [PubMed: 15829077]

Bentin S, Allison T, Puce A, Perez A, McCarthy G. Electrophysiological studies of face perception in humans. Journal of Cognitive Neuroscience. 1996; 8:551–565. doi:10.1162/jocn.1996.8.6.551. [PubMed: 20740065]

Bilalic M, Langner R, Ulrich R, Wolfgang G. Many faces of expertise: Fusiform face area in chess experts and novices. The Journal of Neuroscience. 2011; 31:10206–10214. doi:10.1523/JNEUROSCI.5727-10.2011. [PubMed: 21752997]

Boggan AL, Bartlett JC, Krawczyk DC. Chess masters show a hallmark of face processing with chess. Journal of Experimental Psychology: General. 2012; 141:37–42. doi:10.1037/a0024236. [PubMed: 21787101]

Boutet I, Gentes-Hawn A, Chaudhuri A. The influence of attention on holistic face encoding. Cognition. 2002; 84:321–341. http://dx.doi.org/10.1016/S0010-0277(02)00072-0. [PubMed: 12044738]

Brants M, Wagemans J, Op de Beeck HP. Activation of fusiform face area by greebles is related to face similarity but not expertise. Journal of Cognitive Neuroscience. 2011; 23(12):3949–3958. doi:10.1162/jocn_a_00072. [PubMed: 21671738]

Bukach CM, Bub DN, Gauthier I, Tarr MJ. Perceptual expertise effects are not all or none: Spatially limited perceptual expertise for faces in a case of prosopagnosia. Journal of Cognitive Neuroscience. 2006; 18:48–63. doi:10.1162/089892906775250094. [PubMed: 16417682]

Bukach CM, Cottle J, Ubiwa J, Miller J. Individuation experience predicts other-race effects in holistic processing for both Caucasian and Black participants. Cognition. 2012; 123:319–324. doi:10.1016/j.cognition.2012.02.007. [PubMed: 22397819]

Bukach CM, Gauthier I, Tarr MJ. Beyond faces and modularity: The power of an expertise framework. Trends in cognitive sciences. 2006; 10:159–166. http://dx.doi.org/10.1016/j.tics.2006.02.004. [PubMed: 16516534]

Bukach CM, Philips WS, Gauthier I. Limits of generalization between categories and implications for theories of category specificity. Attention, Perception & Psychophysics. 2010; 72:1865–1874. doi:10.3758/APP.72.7.1865.

Bushnell IW, Sai F, Mullin JT. Neonatal recognition of the mother's face. British Journal of Developmental Psychology. 1989; 7:3–15. doi: 10.1111/j.2044-835X.1989.tb00784.x.

Butko, NJ.; Fasel, I.; Movellan, JR. Learning about humans during the first 6 minutes of life; International Conference on Development and Learning; Indiana. 2006;

Carey S, Diamond R. Are faces perceived as configurations more by adults than by children. Visual Cognition. 1994; 1:253–274. doi: 10.1080/13506289408402302.

Cassia VM, Picozzi M, Kuefner D, Bricolo E, Turati C. Holistic processing of faces and cars in preschool -aged children and adults: Evidence from the composite effect. Developmental Science. 2009; 12:236–248. doi: 10.1111/j.1467-7687.2008.00765.x. [PubMed: 19143797]

Cheung OS, Gauthier I. Selective interference of faces in working memory. on the holistic processin. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36:448–461. doi: 10.1037/a0016471. [PubMed: 20364929]

Cheung OS, Richler JJ, Palmeri TJ, Gauthier I. Revisiting the role of spatia l frequencies in the holistic processing of faces. Journal of Experimental Psychology: Human Perception and Performance. 2008; 34:1327–1336. doi: 10.1037/a0011752. [PubMed: 19045978]

Cheung OS, Richler JJ, Phillips WS, Gauthier I. Does temporal integration of face parts reflect holistic processing? Psychonomic Bulletin & Review. 2011; 18:476–483. doi: 10.3758/s13423-011-0051-7. [PubMed: 21327378]

Chua K-W, Richler JJ, Gauthier I. Becoming a Lunari or Taiyo expert: Learned attention to parts drives holistic processing of faces. Journal of Experimental Psychology: Human Perception and Performance. in press. doi:10.1037/a0035895.

Cochran WG. The combination of estimates from different experiments. Biometrics. 1954; 10:101–129. doi:10.2307/3001666.

Coory MD. Comment on: heterogeneity in meta-analysis should be expected and appropriately quantified. International Journal of Epidemiology. 2010; 39:932. doi: 10.1093/ije/dyp157. [PubMed: 19349478]

Crookes K, McKone E. Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face space. Cognition. 2009; 111:219–247. doi:10.1016/j.cognition.2009.02.004. [PubMed: 19296930]

Curby KM, Gauthier I. A visual short short-term memory advantage for faces. Psychonomic Bulletin & Review. 2007; 14:620–628. doi: 10.3758/BF03196811. [PubMed: 17972723]

Curby KM, Gauthier I. The temporal advantage for individuating objects of expertise: Perceptual expertise is an early riser. Journal of Vision. 2009; 9:1–13. doi:10.1167/9.6.7. [PubMed: 19761298]

Curby KM, Glazek K, Gauthier I. A visual short-term memory advantage for objects of expertise. Journal of Experimental Psychology: Human Perception and Performance. 2009; 35:94–107. doi: 10.1037/0096-1523.35.1.94. [PubMed: 19170473]

Curby KM, Goldstein R, Blacker K. Disrupting perceptua face parts impairs holistic face processing. l grouping o. Attention, Perception & Psychophysics. 2013; 75:83–91. doi:10.3758/s13414-012-0386-9.

Curby KM, Johnson KJ, Tyson A. Face to face with emotion: Holistic face processing is modulated by emotional state. Cognition & Emotion. 2012; 26:93–102. Doi: 10.1080/02699931.2011.555752. [PubMed: 21557121]

Dailey MN, Cottrell GW. Organization of face and object recognition in modular neural networks. Neural Networks. 1999; 12:1053–1074. http://dx.doi.org/10.1016/S0893-6080(99)00050-7. [PubMed: 12662645]

DeGutis J, DeNicola C, Zink T, McGlinchey R, Milberg W. Training with own-race faces can improve processing of other-race faces: Evidence from developmental prosopagnosia. Neuropsychologia. 2011; 49:2505–2513. doi: 10.1016/j.neuropsychologia.2011.04.031. [PubMed: 21570991]

DeGutis J, Wilmer J, Mercado RJ, Cohan S. Using regression to measure holistic face processing reveals a strong link with face recognition ability. Cognition. 2013; 126:87–100. doi: 10.1016/j.cognition.2012.09.004. [PubMed: 23084178]

Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, Dehaene-Lambertz G, Kolinsky R, Morais J, Cohen L. How learning to read changes the cortical networks for vision and language. Science. 2010; 330:1359–1364. doi:10.1126/science.1194140. [PubMed: 21071632]

Derryberry D, Reed MA. Anxiety and attentional focusing: Trait, state, and hemispheric influences. Personality and Individual Differences. 1998; 25:745–761. http://dx.doi.org/10.1016/S0191-8869(98)00117-2.

Dundas EM, Plaut DC, Behrmann M. The Joint Development of Hemispheric Lateralization for Words and Faces. Journal of Experimental Psychology: General. 2013 doi: 10.1037/a0029503.

Eimer M. Effects of face inversion on the structural encoding and recognition of faces: Evidence from event-related brain potentials. Cognitive Brain Research. 2000; 10:145–158. http://dx.doi.org/10.1016/S0926-6410(00)00038-0. [PubMed: 10978702]

Ellis, Paul D. The Essential Guide to Effect Sizes: An Introduction to Statistical Power, Meta-Analysis and the Interpretation of Research Results. Cambridge University Press; United Kingdom: 2010.

Farah MJ, Wilson KD, Drain M, Tanaka JW. What is "special" about face perception. Psychological Review. 1998; 105:482–498. doi: 10.1037/0033-295X.105.3.482. [PubMed: 9697428]

Freire A, Lee K. Face recognition in 4- to 7-year olds: Processing of configural, featural, and paraphernalia information. Journal of Experimental Child Psychology. 2001; 80:347–371. http://dx.doi.org/10.1006/jecp.2001.2639. [PubMed: 11689035]

Gao Z, Flevaris AV, Robertson LC, Bentin S. Priming global and local processing of composite faces: Revisiting the processing-bias effect on face perception. Attention, Perception and Psycho physics. 2011; 73:1477–1486. doi:10.3758/s13414-011-0109-7.

Gauthier I, Bukach C. Should we reject the expertise hypothesis. Cognition. 2007; 103:322–330. http://dx.doi.org/10.1016/j.cognition.2006.05.003. [PubMed: 16780825]

Gauthier I, Klaiman C, Schultz RT. Face composite effects reveal abnormal face processing in Autism Spectrum Disorders. Vision Research. 2009; 49:470–478. doi: 10.1016/j.visres.2008.12.007. [PubMed: 19135077]

Gauthier I, Skudlarski P, Gore JC, Anderson AW. Expertise for cars and birds recruits brain areas involved in face recognition. Nature Neuroscience. 2000; 3:191–197. doi:10.1038/72140.

Gauthier I, Tarr MJ. Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. Journal of Experimental Psychology: Human Perception and Performance. 2002; 28:431–446. doi:10.1037/0096-1523.28.2.431. [PubMed: 11999864]

Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nature Neuroscience. 1999; 2:568–573. doi:10.1038/9224.

Gauthier I, Williams P, Tarr MJ, Tanaka JW. Training "Greeble" experts: A framework for studying expert object recognition processes. Vision Research. 1998; 38:2401–2428. http://dx.doi.org/10.1016/S0042-6989(97)00442-2. [PubMed: 9798007]

Gauthier I, Curran T, Curby KM, Collins D. Perceptual interference supports a non-modular account of face processing. Nature Neuroscience. 2003; 6:428–432. doi:10.1038/nn1029.

Goffaux V. Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. Vision Research. 2009; 49:774–781. doi:10.1016/j.visres.2009.02.009. [PubMed: 19236892]

Goffaux V, Rossion B. Faces are "spatial" – Holistic face perception is supported by low spatial frequencies. Journal of Experimental Psychology: Human Perception and Performance. 2006; 32:1023–1039. doi:10.1037/0096-1523.32.4.1023. [PubMed: 16846295]

Goffaux V, Rossion B. Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. Journal of Experimental Psychology: Human Perception and Performance. 2007; 33:995–1002. doi:10.1037/0096-1523.33.4.995. [PubMed: 17683242]

Green, DM.; Swets, JA. Signal detection theory and psychophysics. Wiley; New York: 1966.

Harrison SA, Gauthier I, Hayward WG, Richler JJ. Other-race effects are quantitative not qualitative in the composite paradigm. in revision.

Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. International Journal of Epidemiology. 2008; 37:1158–1160. doi:10.1093/ije/dyn204. [PubMed: 18832388]

Hills PJ, Cooper RE, Pake JM. First fixations in face processing: The more diagnostic they are the smaller the face-inversion effect. Acta Psychologica. 2013; 142:211–219. doi: 10.1016/j.actpsy.2012.11.013. [PubMed: 23348201]

Hole GJ. Configurational factors in the perception of unfamiliar faces. Perception. 1994; 23:65–74. doi:10.1068/p230065. [PubMed: 7936977]

Humphreys LG. Erroneous interpretation of difference scores: Application to a recent example. Intelligence. 1990; 14:231–233. http://dx.doi.org/10.1016/0160-2896(90)90005-E.

Huntsinger JR. Does positive affect broaden and negative affect nar attentional scope? A new answer to an old question. ro. Journal of Experimental Psychology: General. 2012; 141:595–600. doi: 10.1037/a0027709. [PubMed: 22409665]

Hsiao JH, Cottrell GW. Not all visual expertise is holistic, but it may be leftist: The case of Chinese character recognition. Psychological Science. 2009; 20:455–463. doi:10.1111/j. 1467-9280.2009.02315.x. [PubMed: 19399974]

Jacques C, Rossion B. The initial representation of individual faces in the right occipito-temporal cortex is holistic: Electrophysiological evidence from the composite face illusion. Journal of Vision. 2009; 9:1–16. doi: 10.1167/9.6.8. [PubMed: 19761299]

Johnson MH, Dziurawiec S, Ellis H, Morton J. Newborns' preferential tracking of face-like stimuli and its subsequent decline. Cognition. 1991; 40:1–19. http://dx.doi.org/10.1016/0010-0277(91)90045-6. [PubMed: 1786670]

Kanwisher N, Tong F, Nakayama K. The effect of face inversion on the human fusiform face area. Cognition. 1998; 68:B1–B11. http://dx.doi.org/10.1016/S0010-0277(98)00035-3. [PubMed: 9775518]

Konar Y, Bennett PJ, Sekuler AB. Holistic processing is not correlated with face-identification accuracy. Psychological Science. 2010; 21:38–43. doi:10.1177/0956797609356508. [PubMed: 20424020]

Le Grand R, Cooper PA, Mondloch CJ, Lewis TL, Sagiv N, de Gelder B, Maurer D. What aspects of face processing are impaired in developmental prosopagnosia. Brain and Cognition. 2006; 61:139–158. http://dx.doi.org/10.1016/j.bandc.2005.11.005. [PubMed: 16466839]

Leder H, Bruce V. When inverted faces are recognized: The role of configural information in face recognition. The Quarterly Journal of Experimental Psychology. 2000; 53:513–536. doi: 10.1080/713755889. [PubMed: 10881616]

Loftus GR, Oberg MA, Dillon AM. Linear theory, dimensional theory, and the face-inversion effect. Psychological Review. 2004; 111:835–862. doi: 10.1037/0033-295X.111.4.835. [PubMed: 15482064]

Maurer D, Le Grand R, Mondloch CJ. The many faces of configural processing. Trends in Cognitive Sciences. 2002; 6:255–260. http://dx.doi.org/10.1016/S1364-6613(02)01903-4. [PubMed: 12039607]

Maurer D, Young RE. Newborn's following of natural and distorted arrangements of facial features. Infant Behavior and Development. 1983; 6:127–131. http://dx.doi.org/10.1016/S0163-6383(83)80018-6.

McCandliss BD, Cohen L, Dehaene S. The visual word form area: Expertise for reading in the fusiform gyrus. Trends in Cognitive Sciences. 2003; 7:293–299. http://dx.doi.org/10.1016/S1364-6613(03)00134-7. [PubMed: 12860187]

McGugin RW, Gauthier I. Perceptual expertise with objects predicts another hallmark of face perception. Journal of Vision. 2009; 10:1–12. doi:10.1167/10.4.15. [PubMed: 20465319]

McGugin RW, Gatenby JC, Gore JC, Gauthier I. High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. Proceedings of the National Academy of Sciences. 2012; 109(42):17063–17068. doi: 10.1073/pnas.1116333109.

McGugin RW, Richler JJ, Herzmann G, Speegle M, Gauthier I. The Vanderbilt expertise test reveals domain-general and domain-specific sex effects in object recognition. Vision Research. 2012; 69:10–22. doi:10.1016/j.visres.2012.07.014. [PubMed: 22877929]

McKone E. Configural processing and face viewpoint. Journal of Experimental Psychology: Human Perception and Performance. 2008; 34:310–327. doi:10.1037/0096-1523.34.2.310. [PubMed: 18377173]

McKone E. Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. Vision Research. 2009; 49:268–283. http://dx.doi.org/10.1016/j.visres.2008.10.020. [PubMed: 19022276]

McKone E, Kanwisher N, Duchaine BC. Can generic expertise explain special processing for faces. Trends in Cognitive Sciences. 2007; 11:8–15. http://dx.doi.org/10.1016/j.tics.2006.11.002. [PubMed: 17129746]

McKone E, Yovel G. Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. Psychonomic Bulletin & Review. 2009; 16(5):778–797. doi: 10.3758/PBR.16.5.778. [PubMed: 19815781]

Meissner CA, Brigham JC. Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. Psychology, Public Policy, and Law. 2001; 7:3–35. doi: 10.1037/1076-8971.7.1.3.

Michel C, Rossion B, Han J, Chung C-S, Caldara R. Holistic processing is finely tuned for faces of our own race. Psychological Science. 2006; 17:608–615. doi: 10.1111/j.1467-9280.2006.01752.x. [PubMed: 16866747]

Mondloch C,J, Le Grand R, Maurer D. Configural face processing develops more slowly than featural face processing. Perception. 2002; 31:553–566. doi:10.1068/p3339. [PubMed: 12044096]

Mondloch CJ, Pathman T, Maurer D, Le Grand R, de Schonen S. The composite face effect in six-year-old children: Evidence of adult-like holistic processing. Visual Cognition. 2007; 15:564–577. doi: 10.1080/13506280600859383.

Morton J, Johnson MH. Conspec and conlern: A two- process theory of infant face recognition. Psychological Review. 1991; 98:164–181. doi: 10.1037/0033-295X.98.2.164. [PubMed: 2047512]

Navon D. Forest before perception trees: The precedence of global features in visual. Cognitive Psychology. 1977; 9:353–383. http://dx.doi.org/10.1016/0010-0285(77)90012-3.

Op de Beeck HP, Baker CI. The neural basis of visual object learning. Trends in Cognitive Science. 2010; 14:22–30. doi: 10.1016/j.tics.2009.11.002.

Palermo R, Willis ML, Rivolta D, McKone E, Wilson CE, Calder AJ. Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. Neuropsychologia. 2011; 49:1226–1235. doi:10.1016/j.neuropsychologia.2011.02.021. [PubMed: 21333662]

Richler JJ, Bukach CM, Gauthier I. Context influences holistic processing of nonface objects in the composite task. Attention, Perception, & Psychophysics. 2009; 71:530–540. doi: 10.3758/APP. 71.3.530.

Richler JJ, Cheung OS, Gauthier I. Beliefs alter holistic face processing...if Response Bias is Not Taken into Account. Journal of Vision. 2011a; 11:1–13. doi: 10.1167/11.13.17.

Richler JJ, Cheung OS, Gauthier I. Holistic processing predicts face recognition. Psychological Science. 2011b; 22:464–471. doi:10.1177/0956797611401753. [PubMed: 21393576]

Richler JJ, Cheung OS, Wong AC-N, Gauthier I. Does response interference contribute to face composite effects. Psychonomic Bulletin & Review. 2009; 16:258–263. doi: 10.3758/PBR. 16.2.258. [PubMed: 19293091]

Richler JJ, Gauthier I. When intuition fails to align with data: A reply to Rossion. Visual Cognition. 2013; 21:254–276. doi:10.1080/13506285.2013.796035.

Richler JJ, Gauthier I, Wenger MJ, Palmeri TJ. Holistic processing of faces: Perceptual and decisional components. Journal of Experimental Psychology: Learning, Memory and Cognition. 2008; 34:328–342. doi: 10.1037/0278-7393.34.2.328.

Richler JJ, Mack ML, Gauthier I, Palmeri TJ. Holistic processing of faces happens at a glance. Vision Research. 2009; 49:2856–2861. doi:10.1016/j.visres.2009.08.025. [PubMed: 19716376]

Richler JJ, Mack ML, Palmeri TJ, Gauthier I. Inverted faces are (eventually) processed holistically. Vision Research. 2011; 51:333–342. doi:10.1016/j.visres.2010.11.014. [PubMed: 21130798]

Richler JJ, Palmeri TJ, Gauthier I. Meanings, mechanisms, and measures of holistic processing. Frontiers in Psychology. 2012; 3 doi:10.3389/fpsyg.2012.00553.

Richler JJ, Tanaka JW, Brown DD, Gauthier I. Why does selective attention to parts fail in face processing. Journal of Experimental Psychology: Learning, Memory and Cognition. 2008; 34:1356–1368. doi: 10.1037/a0013080.

Richler JJ, Wong YK, Gauthier I. Perceptual expertise as a shift from strategic interference to automatic holistic processing. Current Directions in Psychological Science. 2011; 20:129–134. doi: 10.1177/0963721411402472. [PubMed: 21643512]

Riesenhuber M, Jarudi I, Gilad S, Sinha P. Face processing in humans is compatible with a simple shape -based model of vision. Proceedings in Biological Sciences. 2004; 271(Suppl. 6):S448–S450. doi: 10.1098/rsbl.2004.0216.

Ro T, Friggel A, Lavie N. Musical expertise modulates the effects of visual perceptual load. Attention, Perception & Psychophysics. 2009; 71:671–674. doi: 10.3758/APP.71.4.671.

Robbins R, McKone E. No face-like processing for objects-of-expertise in three behavioural tasks. Cognition. 2007; 103:34–79. http://dx.doi.org/10.1016/j.cognition.2006.02.008. [PubMed: 16616910]

Rogosa, D. Myths about longitudinal research. In: Schaie, KW.; Campbell, RT.; Meredith, M.; Rawlings, SC., editors. Methodological Issues in Aging Research. Springer; New York, NY: 1988. p. 171-209.

Rosenthal, R.; Rosnow, RL.; Rubin, DB. Contrasts and effect sizes in behavioral research: A correlational approach. Cambridge University Press; New York, NY: 2000. http://dx.doi.org/10.1016/j.beth.2006.05.002

Rossion B. Picture-plane inversion leads to qualitative changes of face perception. Acta Psychologia. 2008; 128:274–289. doi:10.1016/j.actpsy.2008.02.003.

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

Rossion B, Boremanse A. Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. Journal of Vision. 2008; 8:1–13. doi: 10.1167/8.4.3. [PubMed: 18484842]

Rossion B, Gauthier I, Tarr MJ, Despland PA, Bruyer R, Linotte S, Crommelinck M. The N170 occipito-temporal component is enhanced and delayed to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. Neuroreport. 2000; 11:69–74. [PubMed: 10683832]

Sagiv N, Bentin S. Structural encoding of human and schematic faces: Holistic and part-based processes. Journal of Cognitive Neuroscience. 2001; 13:937–951. doi: 10.1162/089892901753165854. [PubMed: 11595097]

Schimmack U. The ironic effect of significant results on the credibility of multiple-study articles. Psychological methods. 2012; 17:551. doi:10.1037/a0029487. [PubMed: 22924598]

Schultz RT. Developmental deficits in social perception in Autism: The role of the amygdala and fusiform face area. International Journal of Developmental Neuroscience. 2005; 23:125–141. http://dx.doi.org/10.1016/j.ijdevneu.2004.12.012. [PubMed: 15749240]

Schwartz BL, Marvel CL, Drapalski A, Rosee RB, Deutsch SI. Configural processing in face recognition in schizophrenia. Cognitive Neuropsychiatry. 2002; 7:15–39. doi: 10.1080/13546800143000113. [PubMed: 16528403]

Searcy JH, Bartlett JC. Inversion and processing of component and spatial-relational information in faces. Journal of Experimental Psychology: Human Perception and Performance. 1996; 22:904–915. doi:10.1037/0096-1523.22.4.904. [PubMed: 8756958]

Sekuler AB, Gaspar CM, Gold JM, Bennett PJ. Inversion leads to quantitative, not qualitative, changes in face processing. Current Biology. 2004; 14:391–396. http://dx.doi.org/10.1016/j.cub.2004.02.028. [PubMed: 15028214]

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

Tanaka JW, Farah MJ. Parts and wholes in face recognition. The Quarterly Journal of Experimental Psycholgy. 1993; 46A:225–245. doi: 10.1080/14640749308401045.

Tanaka JW, Kaiser MD, Butler S, Le Grand R. Mixed emotions: Holistic and analytic perception of facial expressions. Cognition & Emotion. 2012; 26(6):961–977. doi: 10.1080/02699931.2011.630933. [PubMed: 22273429]

Thomas DR, Zumbo BD. Difference Scores From the Point of View of Reliability and Repeated-Measures ANOVA In Defense of Difference Scores for Data Analysis. Educational and Psychological Measurement. 2012; 72:37–43. doi: 10.1177/0013164411409929.

Todorov A, Loehr V, Oosterhof NN. The obligatory nature of holistic processing of faces in social judgments. Perception. 2010; 39:514–532. doi:10.1068/p6501. [PubMed: 20514999]

Tsao DY, Livingstone MS. Mechanisms of face perception. Annual Review of Neuroscience. 2008; 31:411–437. doi:10.1146/annurev.neuro.30.051606.09423.

Valentine T, Bruce V. Mental rotation of faces. Memory and Cognition. 1988; 16:556–566. doi: 10.3758/BF03197057. [PubMed: 3193887]

Wang R, Li J, Fang H, Tian M, Liu J. Individual differences in holistic processing predict face recognition ability. Psychological Science. 2012; 23:169–177. doi: 10.1177/0956797611420575. [PubMed: 22222218]

Weiner KS, Grill-Spector K. Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. Psychological Research. 2013; 77(1): 74–97. doi: 10.1007/s00426-011-0392-x. [PubMed: 22139022]

Wilhelm O, Herzmann G, Kunina O, Danthiir V, Schact A, Sommer W. Individual differences in perceiving and recognizing faces – one element of social cognition. Journal of Personality and Social Psychology. 2010; 99:530–548. doi: 10.1037/a0019972. [PubMed: 20677889]

Willenbockel V, Fiset D, Chauvin A, Blais C, Arguin M, Tanaka JW, Bub DN. Does face inversion change spatial frequency tuning. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36:122–135. doi: 10.1037/a0016465. [PubMed: 20121299]

Wong AC-N, Bukach CM, Hsiao J, Greenspon E, Ahern E, Duan Y, Lui KFH. Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. Journal of Vision. 2012; 12:1–15. doi:10.1167/12.13.7.

Wong AC-N, Bukach CM, Yuen C, Yang L, Leung S, Greenspon E. Holistic processing of words modulated by reading experience. PLoS ONE. 2011; 6:e20753. doi: 10.1371/journal.pone. 0020753. [PubMed: 21698240]

Wong YK, Folstein JR, Gauthier I. The Nature of Experience Determines Object Representations in the Visual System. Journal of Experimental Psychology: General. 2012; 141:682–698. doi: 10.1037/a0027822. [PubMed: 22468668]

Wong AC-N, Gauthier I. An analysis of letter expertise in a levels-of-categorization framework. Visual Cognition. 2007; 15:854–879. DOI:10.1080/13506280600948350.

Wong YK, Gauthier I. Holistic processing of musical notation: Dissociating failures of attention in experts and novices. Journal of Cognitive and Affective Behavioral Neuroscience. 2010; 10:541–551. doi:10.3758/CABN.10.4.541.

Wong AC-N, Palmeri TJ, Gauthier I. Conditions for face-like expertise with objects: Becoming a Ziggerin expert – but which type? Psychological Science. 2009; 20:1108–1117. doi: 10.1111/j. 1467-9280.2009.02430.x. [PubMed: 19694980]

Wong AC-N, Palmeri TJ, Rogers BP, Gore JC, Gauthier I. Beyond shape: how you learn about objects affects how they are represented in visual cortex. PLoS ONE. 2009; 4:e8405. doi: 10.1371/journal.pone.0008405. [PubMed: 20027229]

Xu Y. Revisiting the role of the fusiform face area in visual expertis. Cerebral Cortex. 2005; 15:1234–1242. doi: 10.1093/cercor/bhi006. [PubMed: 15677350]

Xu B, Tanaka JW. Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. Journal of Vision. 2013; 13:1–16. doi: 10.1167/13.2.22.

Yin RK. Looking at upside-down faces. Journal of Experimental Psychology. 1969; 81:141–145. http://dx.doi.org.proxy.library.vanderbilt.edu/10.1037/h0027474.

Young AW, Hellawell D, Hay DC. Configurational information in face perception. Perception. 1987; 16:747–759. doi:10.1068/p160747. [PubMed: 3454432]

Yovel G, Kanwisher N. The neural basis of the behavioral face-inversion effect. Current Biology. 2005; 15:2256–2262. http://dx.doi.org/10.1016/j.cub.2005.10.072. [PubMed: 16360687]

Zhou G, Cheng Z, Zhang X, Wong ACN. Smaller holistic processing of faces associated with face drawing experience. Psychonomic Bulletin & Review. 2012; 19(2):157–162. doi: 10.3758/s13423-011-0174-x. [PubMed: 22215464]

Zhu Q, Song Y, Hu S, Li X, Tian M, Zhen Z, Dong Q, Kanwisher N, Liu J. Heritability of the specific cognitive ability of face perception. Current Biology. 2010; 20:1–6. doi: 10.1016/j.cub. 2009.11.067. [PubMed: 20036540]
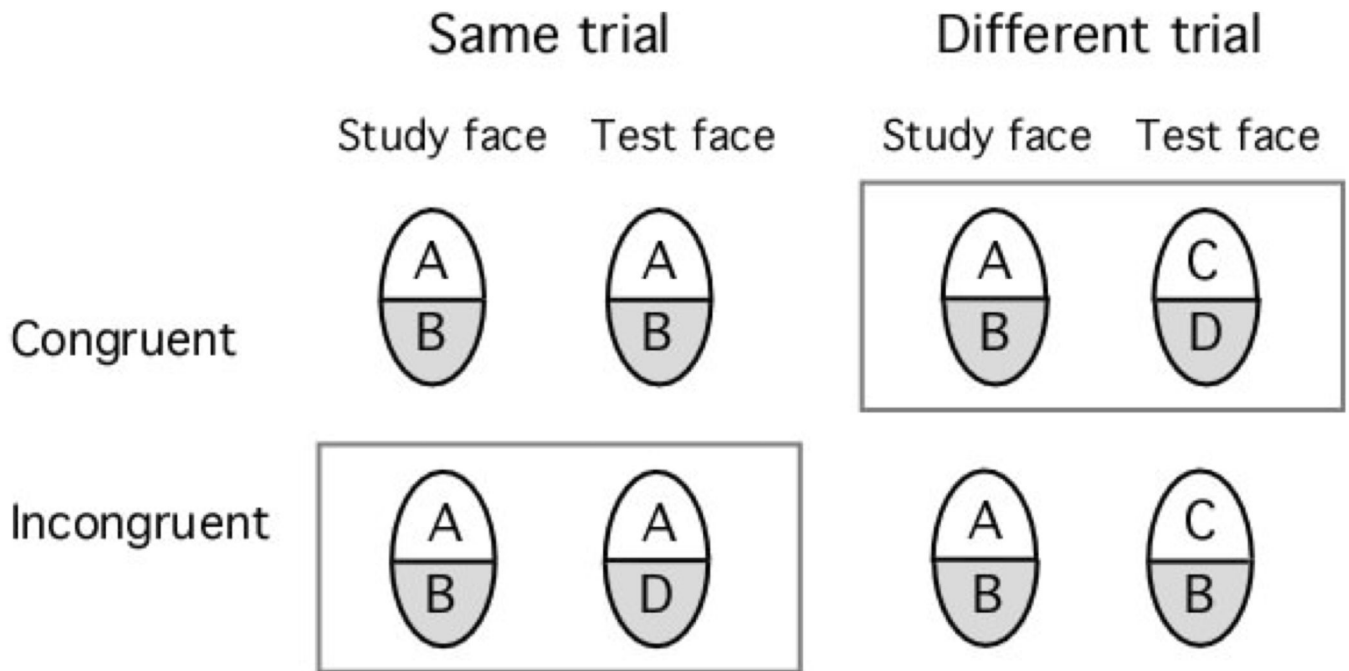
Figure 1.
Design of the composite task for aligned trials. Letters represent facial identities. Task-relevant (target) face halves are shown in white. Task-irrelevant (distractor) halves are shown in gray. In *same* trials, target halves of the study and *different* test faces are the same; in *same* trials, target halves are different. In congruent trials, the irrelevant halves of the study and test faces are the same; in incongruent *same* trials, the irrelevant halves are different. In congruent *different* trials, the irrelevant halves of the study and test faces are different; in incongruent *different* trials, the irrelevant halves are the same. In the partial-design version of this task, only the trial types outlined in the gray boxes are presented. In the complete design, all trial types are presented. In both designs face halves can be aligned or misaligned.
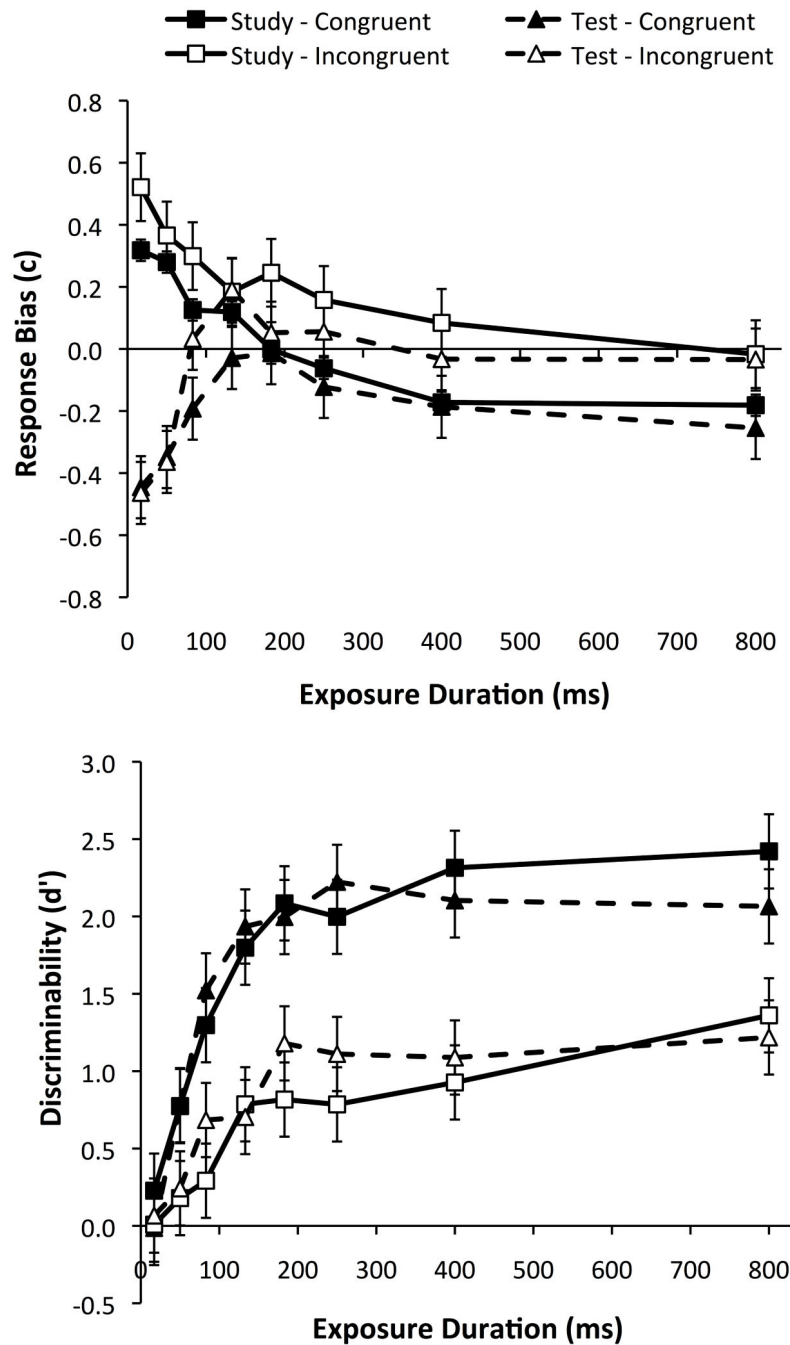
Figure 2.
Response bias (top panel) and discriminability (bottom panel) from Richler, Mack, et al. (2009). Congruent trials are shown as black markers, incongruent trials are shown as white markers. Squares indicate conditions where presentation times for study faces were manipulated, and triangles indicate conditions where presentations times of test faces were manipulated. For rapid exposure durations, response bias is sensitive to whether presentation time for the study or test face is manipulated. Whether presentation time for the study or test face is manipulated does not influence discriminability.
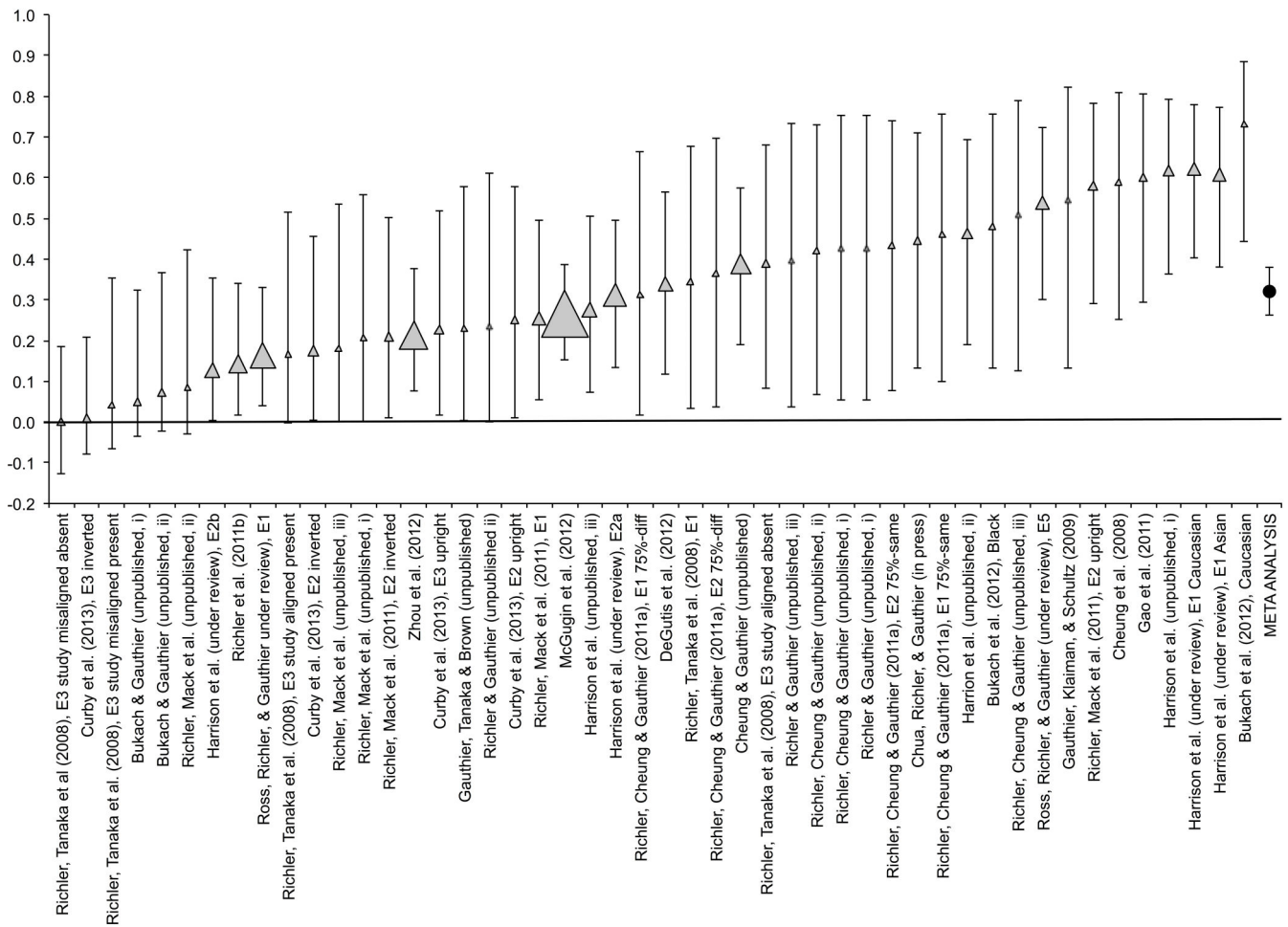
Figure 3.
Effect size (η^2_p) of holistic processing in the complete design (congruency × alignment interaction in d'). The size of the triangle markers are scaled by sample size. The black circle is the mean weighted effect size from the meta-analysis. Error bars show 95% confidence intervals.
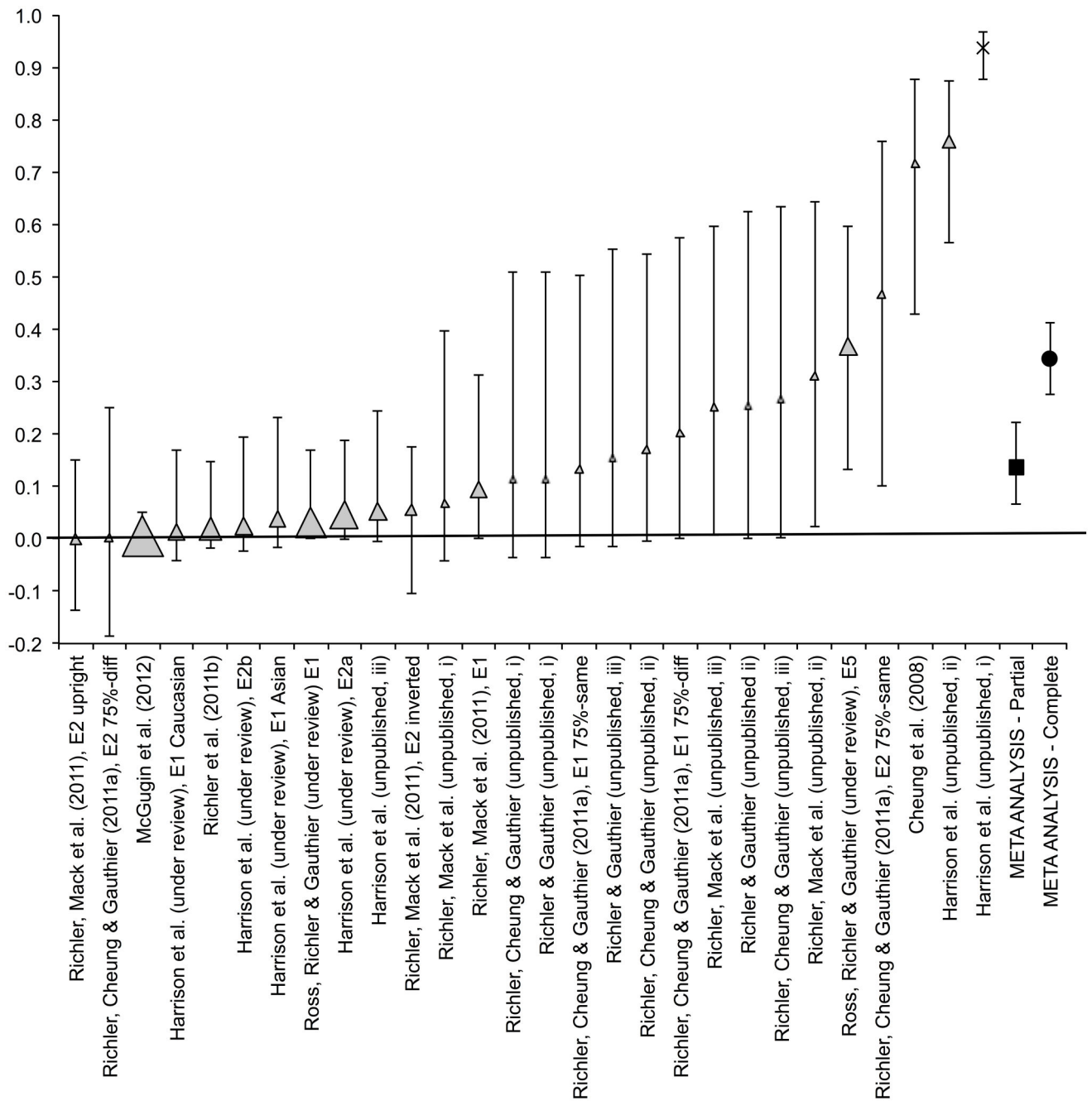
Figure 4.
Effect size (η^2_p) of holistic in the partial design (alignment effect in hit rate on same-incongruent trials) calculated from a subset of trials run in the context of the complete design. The size of the triangle markers are scaled for sample size. The black square is the mean weighted effect size for holistic processing in the partial design from the meta-analysis. The black circle is the mean weighted effect size for holistic processing in the complete design for this subset of studies. Error bars show 95% confidence intervals
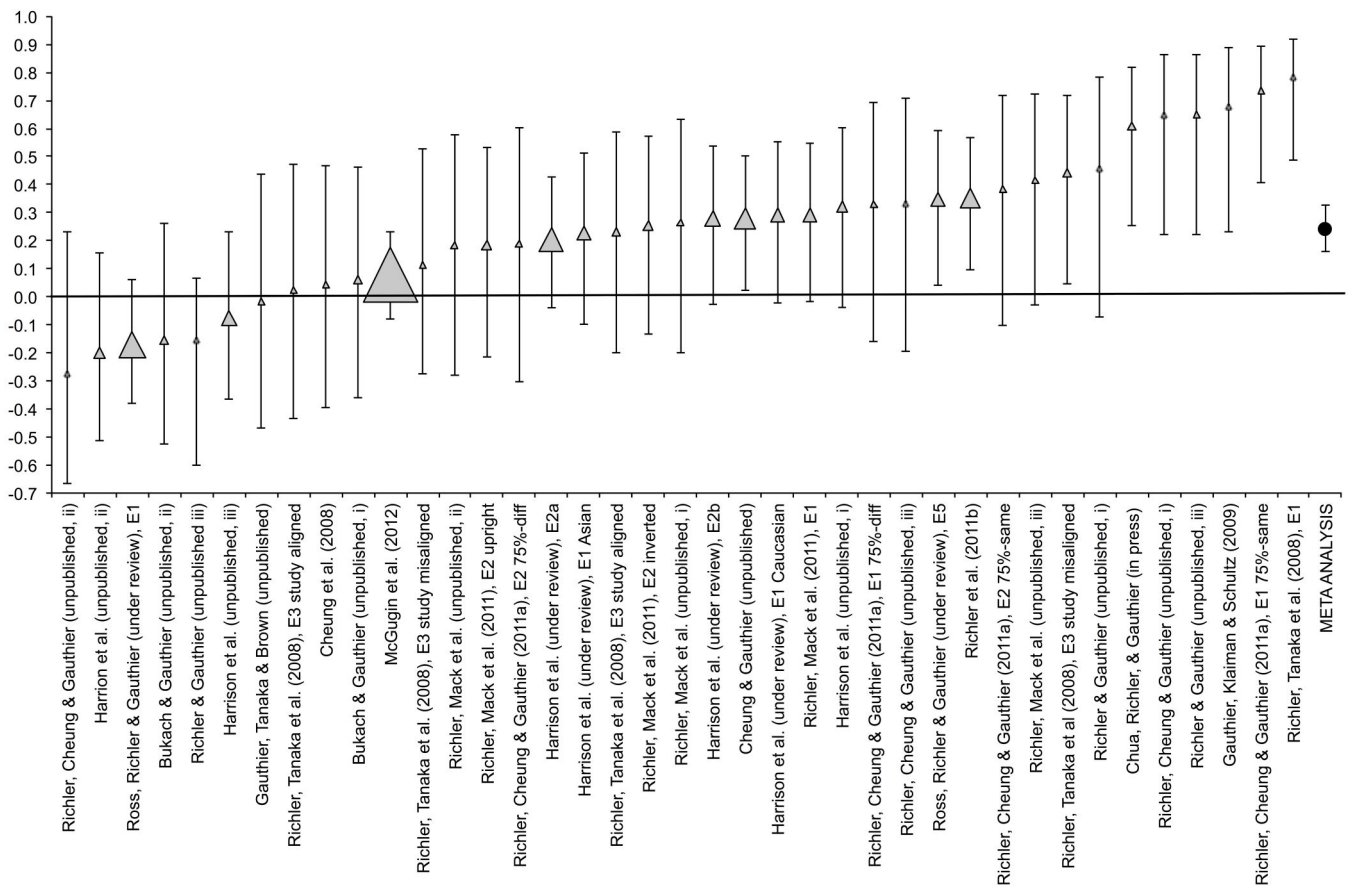
Figure 5.
Effect size of the correlation (*r*) between the congruency effect-aligned and congruency effect-misaligned. The size of the triangle markers are scaled for sample size. The black circle shows the mean-weighted effect size. Error bars show 95% confidence intervals.
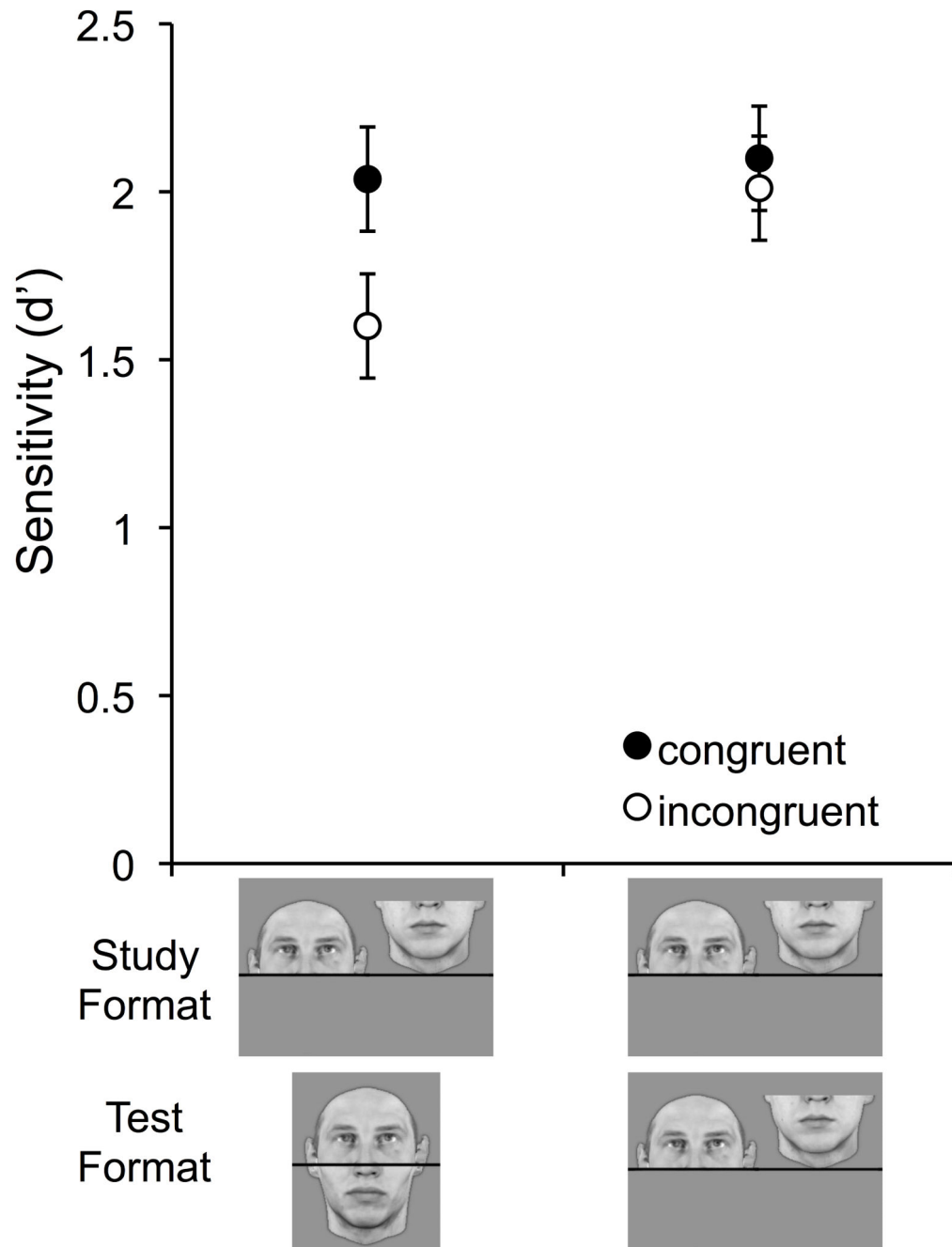
Figure 6.
Re-plotted data from Richler, Tanaka, Brown & Gauthier (2008), Experiment 2. In this
version of the composite task, the study face was always presented in an adjacent format,
where the top and bottom face halves were presented side-by-side. The test face could be
either aligned or adjacent. A congruency effect was observed when the test face was aligned,
even though the study face was presented in an unusual configuration that could not be
encoded to fit a face template.

Table 1

Experiments included in the meta-analysis of holistic processing of faces. Unless otherwise noted, faces were upright Caucasian faces, the study face was aligned on all trials, and when both top and bottom face halves were tested the target half was randomized

Paper	Exp	Condition	Tested Part	Study Face	Test Face (ms)
^ Bukach et al. (2012) [a]		Caucasian Ss	top & bottom	1500ms misaligned on misaligned test trials	until response
		Black Ss	top & bottom	1500ms misaligned on misaligned test trials	until response
Bukach & Gauthier	i		top & bottom		until response
(unpublished data)	ii		top & bottom	misaligned	until response
Cheung & Gauthier (unpublished data)			top or bottom (between Ss)	800ms	800ms
* Cheung et al. (2008)			top	600ms	until response
Chua, Richler, & Gauthier (in press)		trained Alien faces	top & bottom (blocked)	200ms	200ms
^ Curby et al. (2013)	2		top & bottom	1000ms	until response
	2	inverted faces	top & bottom	1000ms	until response
	3		top & bottom	1000ms	200ms
	3	inverted faces	top & bottom	1000ms	200ms
^ DeGutis et al. (2013)			top	400ms	400ms
^ Gao et al. (2011) [b]			top	250ms	until response
^ Gauthier et al. (2009)		control group	top & bottom (blocked)	700ms	until response
Gauthier, Tanaka & Brown (unpublished data)			top & bottom (blocked)	700ms	until response
*Harrison, Gauthier,	i	Caucasian Ss	top	600ms	until response
Hayward & Richler (unpublished data)		Asian & Caucasian faces		misaligned on misaligned test trials	
	ii	Asian Ss Asian & Caucasian faces	top	600ms misaligned on misaligned test trials	until response
	iii	In-group/out-group own-race faces	top	600ms	until response
*Harrison et al. (under review)	1	Caucasian Ss Asian & Caucasian faces	top	800ms	50/183/800ms
	1	Asian Ss Asian & Caucasian faces	top	800ms	50/183/800ms
	2a	in-group/out-group own-race faces	top	600ms	33/83/400ms
	2b	in-group/out-group own race faces	top	600ms	until response

Paper	Exp	Condition	Tested Part	Study Face	Test Face (ms)
* McGugin et al. (2012) [c]	3		top	200ms	200ms
*Richler, Cheung &	i		top & bottom	600ms	until response
Gauthier	ii	instructed that	top & bottom	600ms	until response
(unpublished data)		distractor is "same" on 75% of trials			
	iii	instructed that distractor is "different" on 75% of trials	top & bottom	600ms	until response
* Richler, Cheung & Gauthier (2011a)	1	instructed that target is "same" on 75% of trials	top & bottom	600ms	until response
	1	instructed that target is "different" on 75% of trials	top & bottom	600ms	until response
	2	target "same" on 75% of trials (no instructions)	top & bottom	600ms	until response
	2	target "different" on 75% of trials (no instructions)	top & bottom	600ms	until response
* Richler, Cheung & Gauthier (2011b) [c]			top	200ms	200ms
*Richler & Gauthier (unpublished data)	i		top & bottom	600ms	until response
	ii	standing	top & bottom	600ms	until response
	iii	sitting & standing	top & bottom	600ms	until response
*Richler, Mack, Palmeri & Gauthier (unpublished data)	i	inverted faces	top & bottom	800ms	until response
	ii	inverted faces	top & bottom	800ms, misaligned	until response
	iii		top & bottom	800ms	until response
* Richler, Mack et al. (2011)	1	upright & inverted faces	top & bottom	800ms	until response
	2		top & bottom	800ms	50/183/800ms
	2	inverted faces	top & bottom	800ms	50/183/800ms
Richler, Tanaka et al. (2008) [d]	1		top & bottom	700ms	until response
				aligned & misaligned	
	3	test face absent	top & bottom	800ms	800ms
	3	test face absent	top & bottom	800ms, misaligned	800ms
	3	test face present	top & bottom	800ms	until response
	3	test face present	top & bottom	800ms, misaligned	until response
*Ross, Richler, & Gauthier (under review)	1		top	200ms	200ms
	5		top	200ms	200ms
^ Zhou et al. (2012)		Asian Ss, Asian faces	top	200ms	until response

[a]Analysis of d' reported in supplemental material.

[b]Performance measured in Az.

[c]Only data for Caucasian participants are reported in the published paper. Data from all participants were included in the meta-analysis.

[d]Published data are reported in Az, but d' was used for the meta-analysis.

[*]Included in meta-analysis of partial design HP

[^]Not included in meta-analysis of the correlation between congruency effect-aligned and congruency effect-misaligned.

Table 2

Provisory conclusions of the literature review of holistic processing (HP) based only on the complete design of the composite task

Topic	Conclusion	Further questions/limitations
HP in non-face objects	HP is obtained with several categories of non-face objects given expertise.	How does HP for faces and objects relate in terms of individual differences?
Mechanisms of HP	Representational mechanisms may exist, but attentional mechanisms could be sufficient; HP of objects correlates with activity in FFA.	Studies aimed at supporting representational account using the complete design are needed. HP of faces in complete design should be measured against activity in FFA for faces.
HP and competition	HP may be the bottleneck of expert processing.	The mechanisms underlying competition need to be modeled explicitly.
HP vs. strategic congruency effects	HP is not susceptible to strategic or contextual manipulations that influence failures of selective attention in novices.	Are top-down effects observed in novices antecedents of automatic effects observed in experts?
HP and prosopagnosia	There is too little data from the complete design in acquired prosopagnosia, a highly heterogeneous population, to draw firm conclusions.	HP in developmental (congenital) prosopagnosia needs to be assessed with the complete design.
HP and autism	More data from complete design needed in this population. Results so far suggest processing of faces is similar to that of objects in novices.	Individuals with autism need to be tested using an experimental design that does not produce strategically-induced congruency effects.
HP of different kinds of faces	Inverted, other-race, and out-group faces are processed holistically in at least some conditions, even though these faces may lead to poorer recognition.	The temporal dynamics of HP may be important in understanding how performance is limited by these different manipulations.
Many questions about HP have only been investigated with the partial design:	Does HP vary with distance/size? At what age is HP mature? Does HP vary with viewpoint? Does early visual deprivation affect HP? Is HP of identity independent from processing other dimensions such as gaze or facial expression?	