Published in final edited form as:
Vis cogn. 2013 May 10; 21(2): . doi:10.1080/13506285.2013.796035.

When Intuition Fails to Align with Data: A Reply to Rossion (2013)

Jennifer J. Richler* and **Isabel Gauthier**
Vanderbilt University

Abstract

Holistic processing, a hallmark of face perception, is often measured in the so-called composite paradigm, in which participants are asked to match part of a stimulus while ignoring another part. In prior work, we recommended against the use of one version of the composite task we call the partial design, on the basis of confounds with response biases. Rossion wrote a lengthy piece that reviews the work that he has published using this design, raising a large number of criticisms, both about an alternative measure of holistic processing that we have used and advocated (which we call the complete design) and about our work in general. In this reply, we have limited our discussion to those issues that would be relevant to a researcher looking to decide which version of this composite paradigm to use, as we doubt a comprehensive reply would be of significant interest outside a very small circle.

When a critique entitled *100 Authors Against Einstein* was published, Einstein's reply is said to have been "If I were wrong, one would be enough".[1]

In a series of articles (Cheung et al. 2008; Richler, Cheung & Gauthier, 2011a,b; Richler, Mack et al., 2011), we have argued that there is one basic problem with a measure of holistic processing (the partial design, or PD) that has been widely used in the face recognition literature: a response bias is confounded with the critical conditions. It appears that no simple response was sufficient to address this concern such that, while our concerns never targeted his work specifically, Rossion devoted roughly one third of his lengthy review of holistic face processing to criticizing several aspects of our work, including the alternative measure of holistic processing (the complete design, or CD, adapted from Farah et al. 1998) that we use and have advocated, and to attacking several conclusions we have drawn over a number of articles. Rossion's piece makes a large number of claims - we counted at least 168 claims about holistic processing and face perception, with over 60 specifically directed toward us and our work. We have many serious issues with the many claims made by Rossion: some of them are opinions based on intuitions with which we disagree, some of them mischaracterize our work and are demonstrably false, some of them clearly ignore published evidence, and some of them criticize our empirical work on the basis of thought experiments and imagined data.

Although we welcome an explicit defense of the PD as a counterpoint to our work, a great many of the issues raised by Rossion are not directly relevant to the problem facing a researcher in the field of face recognition who is interested in measuring holistic processing

***Corresponding Author:** Jennifer J. Richler jennifer.j.richler@vanderbilt.edu REGULAR MAIL (via U.S. Postal Service) Vanderbilt University PMB 407817 2301 Vanderbilt Place Nashville, TN 37240-7817, USA COURIER MAIL (via Fed Exp, UPS) Department of Psychology 301 Wilson Hall Vanderbilt University Nashville, TN 37240, USA.

[1]We thank Daniel Bub for bringing this anecdote to our attention. It goes without saying that we do not take our insights about the measurement of holistic processing to be on the same plane as Einstein's theories.

and must choose between the PD and CD versions of the composite task. Indeed, while opinions abound, all the work that has directly compared the PD and CD empirically has suggested problems with the PD and concluded that the CD is preferable (DeGutis et al., 2013; Cheung et al., 2008; Richler, Cheung & Gauthier, 2011,a,b; Richler, Mack et al., 2011). In that context, we deplore that some of this published evidence was not discussed by Rossion, and that many conclusions were made on the basis of armchair experiments. Decisions about measurement should not be settled on the basis of insights, preferences, or whether a measure produces results that are aligned with one's favored theory.

Therefore, in preparing this rebuttal, as hard as it was, we chose to say almost nothing about many theoretical issues discussed by Rossion (2013). For instance, on questions of whether holistic processing has a perceptual or a decisional locus, or on whether inverted and upright faces are processed in qualitatively vs. quantitatively different ways, the jury is still out; and, as long as basic questions of measurement are not resolved it will be hard – arguably even impossible – to reach consensus. Instead, we chose to focus on the claims and issues relevant to important choices regarding measurement, a good place to start, as even researchers who disagree about theory must eventually agree on measurement. Because the PD and CD often provide qualitatively different conclusions, the reader is presented with the following options: 1) the PD and CD measure two different and theoretically interesting meanings of holistic processing; 2) only one of the two measures is valid; or 3) both of these measures lack the validity necessary to quantify holistic processing.

In what follows, we directly respond to what we view as the most important claims about measurement made by Rossion, and directly test some of these claims using either new data or re-analysis of existing data. We apologize to the reader who is unlikely to find closure here on many of the dozens of issues raised by Rossion. Brief responses to some of the claims that we judged too peripheral to bear directly on the choice of measurement but that, frankly, we could not let stand uncorrected, can be found in the Appendix.

On perceptual integration vs. a failure of selective attention: “one sees two identical top halves as being different when they are aligned with different bottom halves…The phenomenon is not conceptualized as a failure of selective attention: during the composite face task, an observer is asked to judge the top half of a face only and keep fixation on it”

In his article, Rossion gives us a window into his phenomenology and into the thought process that led him to infer how holistic processing should be measured. He perceives the same top half of a composite differently when it is aligned with different bottom halves. He argues that this perception is the result of perceptual integration, and that there is no evidence that a failure of selective attention is involved because he (like our participants and his) was asked to pay attention to the top part only, and is free to fixate it. Yet, it seems absurd to us not to call this a failure of selective attention, since Rossion was asked to selectively attend, and he clearly fails. Indeed, Rossion *must* be allowing information from the bottom halves to influence his judgment, since there is nothing else to explain why the tops look different.

Perhaps Rossion mistakes his intuition that he cannot control this perception with some special insight into the mechanisms that underlie it? To call the composite task, *any version* of the composite task, a selective attention task, is to recognize that subjects were asked to ignore a part, and to call the resulting behavior *a failure of selective attention* is to recognize that they fail at this.

Despite our best efforts (Richler, Palmeri & Gauthier, 2012), Rossion mistakes our use of the phrase “selective attention” to describe the task for a statement about the mechanisms that underlie holistic processing. Indeed, we have advanced (Richler, Gauthier, Wenger & Palmeri, 2008; Richler, Tanaka, Brown & Gauthier, 2008), revised (Richler, Mack et al., 2011) and refined (Richler, Wong & Gauthier, 2011) our ideas about the locus of holistic processing, apparently confusing Rossion to the point that he complains not to know our position on this issue. Suffice it to say, our position is that holistic processing may have an attentional locus, and that it would be unwise to reject this hypothesis on the basis of Rossion's phenomenological report. Attention includes powerful mechanisms known to act at all levels of the visual system past the retina. Even a powerful illusion with an undeniable phenomenology, such as the McGurk effect - whereby a phoneme is perceived differently because of the visual perception of incongruent moving lips - is reduced by a concurrent attentional load (Alsius et al., 2007). Spatial attention can influence multisensory processing as early as 80 ms in EEG recordings (Talsma & Woldorff, 2005). Likewise, the conscious gestalt grouping of elements into a holistic percept, also known as spatial binding, was shown, using a combination of fMRI and TMS methodologies, to depend on parietal activity associated with attention and selection (Zaretskaya, Anstis & Bartels, 2013). Even the tuning properties of V1 cells have been found to depend on the current task context (McManus, Li & Gilbert, 2011). Therefore, regardless of whether attention is important to the face composite illusion Rossion describes, it should be clear that his phenomenology is not sufficient grounds to reject an attentional contribution.

Rossion further invites readers to share his own phenomenology and intuitions by asking them to consider the composite stimuli “without pressing a key”, as if this could isolate bottom-up processes. Rather, it seems to confuse decisional processes with motor commands. Since the focus of this rebuttal is on how we might choose to measure holistic processing, we will have to ask for some key presses and make sure that we can interpret them. This is what we address next.

On the specificity of CD effects: “In reality, these studies reveal a general interference/congruency effect that can be found with pretty much any kind of visual display made of two congruent or incongruent elements”

Rossion argues that because the CD is a task that measures selective attention, like Stroop or Flanker, we will find evidence for interference that we interpret as holistic processing for “just about anything”. Indeed, according to Rossion, in our paradigm we would find holistic processing of face halves juxtaposed with line drawing of animals (see his Figure 37). Rossion drew this conclusion from data he imagined in an experiment that he did not run. Yet, in reality, there is published evidence (from real, not imagined, experiments) that for novices, objects are actually *not* processed holistically in the CD (e.g., Richler, Mack et al., 2011; Wong et al., 2009a). But perhaps Rossion is onto something about face parts, goats and bunnies (Rossion, 2013, Figure 37)?

Therefore, we decided to address this criticism by running the armchair experiment that Rossion only imagined (see Figure 1). We asked participants to judge whether top face halves were the same or different, while ignoring color line drawings of animals. We ran all the conditions necessary for the CD: the relationship between the target face half and the irrelevant animal could be congruent or incongruent, and the target face half and irrelevant animal could be aligned or misaligned.

Our results are shown in Figure 2. A 2 × 2 repeated-measures ANOVA on sensitivity (d') with congruency (congruent/incongruent) and alignment (aligned/misaligned) as factors

revealed no significant effects (main effect of congruency: $F_{1,18}$ = .28, p = 60, η^2_p .02; main effect of alignment: $F_{1,18}$ = .39, p = .54, η^2_p = .02; congruency × alignment interaction: $F_{1,18}$ < .00, p = .99, η^2_p < .00). Nor were any effects significant in the same analysis of correct RTs (main effect of congruency: $F_{1,18}$ = 2.45, p = .14, η^2_p = .12; main effect of alignment: $F_{1,18}$ = .04, p = .85, η^2_p = .002; congruency × alignment interaction: $F_{1,18}$ = 1.21, p = .29, η^2_p = .06). Note that the direction of the RT effects is opposite to any prediction of holistic processing (i.e., faster RT on incongruent trials). These null results are unlikely due to low power. Besides the fact that the effect sizes are all extremely small, a meta-analysis of more than 25 composite task experiments using the CD and over 40 independent samples indicated that the average effect size (η^2_p) for the congruency × alignment interaction indicative of holistic processing in d' is .31, 95% CI: .27, .35 (Richler & Gauthier, in prep). An *a priori* power analysis indicates that to observe an effect of this size 80% of the time with alpha of .05 requires a sample size of 9 participants. In contrast, assuming it is real, the congruency × alignment effect size obtained in this implementation of Rossion's goat and bunny experiment would require 27,471 participants.

Of course, in hindsight, Rossion's rationale for discussing this armchair experiment may not have been only to make the prediction that anything, including goats and bunnies, would produce a congruency effect in the CD, a prediction which is absurd given the published evidence (e.g., Richler, Mack et al., 2011; Wong et al., 2009a). Rather, Rossion argues that this example illustrates the necessity of the misaligned condition in the CD. Interestingly, what can be appreciated from the results is that when testing a category for which there is little to no evidence of a congruency effect in the aligned condition (the literature shows this is the case for non-face objects in novices for instance, with very small effect sizes of congruency effects for objects compared to those for faces), the misaligned condition is in essence a luxury. In our implementation of Rossion's goat and bunny experiment, the congruency effect is not significant, but more importantly it is much smaller than what we regularly document for faces (e.g., in d' η^2_p =.02 here vs. .51–.66 in Richler, Cheung & Gauthier, 2011a), and, if it is a true effect, would require 137 participants to obtain with 80% power. While in the PD no inference at all can be made without the misaligned condition, in the CD, one can calculate a congruency effect and draw an inference about the absence of holistic processing – for some research questions this may be sufficient. If a congruency effect is detected in the aligned condition, and researchers are studying upright faces in normal observers, there may also be little need to replicate the absence (or reduced) congruency effect in the misaligned condition. But when there is a need to ensure that the congruency effect for a new category, or for a new population, or under new conditions, is of the same flavor as the typical face result, or if researchers are aiming to quantify the magnitude of holistic processing across subjects, then a misaligned condition is useful, and indeed we have advocated it (e.g., Richler, Wong & Gauthier, 2011) and have used it in most of our studies (Cheung et al., 2008; Gauthier et al., 2009; McGugin et al., 2012; Richler, Bukach & Gauthier, 2009; Richler, Cheung & Gauthier, 2011a,b; Richler, Gauthier et al., 2008; Richler, Mack et al., 2011; Richler, Tanaka et al., 2008; Wong et al., 2009a).

On the stimuli used in the CD: "*The most salient problem*"

A great deal of Rossion's defense of the PD is an attack of the CD and we applaud him for realizing that regardless of whether there are bias problems in the PD, perhaps there are entirely different problems with the CD that also invalidate its use. But what he notes as "the most salient problem" is not an issue that is specific at all to the CD, it appears instead to be cautioning the reader against other aspects of our work. Specifically, we create our composite stimuli by randomly combining tops and bottoms of different faces, making sure that a composite is never made out of two parts of the same individual so that a decision on a different trial cannot be made based on the parts not fitting together as well as in the

original. In one case (Cheung et al., 2008) where we used the original stimuli from a paper by Rossion (Goffaux & Rossion, 2006), the stimuli varied enough in width that, when they were randomly combined, Rossion claims we created “misaligned aligned faces” that he claims should minimize holistic processing. The problem is that Rossion only presents one such image (in his Figure 41) as support for his intuition. That Rossion can intuit about holistic processing from a single face offers a critical insight into his conception of holistic processing: it seems to have more to do with how one face part “fits” nicely with another part, rather than how it appears different in different part contexts. The reality is that the latter definition is what both the PD and CD are designed to measure, and it is easy to use Rossion's own face images (from his Figure 1) to show that a top part juxtaposed with three bottoms of different widths “looks different” (see Figure 3).

Like Rossion, in our work we have identified the “most salient problem” with the PD: the PD is susceptible to the influence of spurious response biases that may not be related to holistic processing (e.g., Cheung et al., 2008; Richler, Cheung, & Gauthier, 2011a,b; Richler, Mack et al., 2011). In the next sections, we respond to three of Rossion's claims about the PD that are all related to this single issue of response bias.

On the trials necessary in a composite task experiment: “'same' trials are relevant in this composite face paradigm, while `different' trials are not relevant.”

When we run experiments, even if we are only interested in some trials or some conditions, we need to include others so that participants are not cued to what we are interested in, or in how they “should” respond. Rossion argues that the different trials in the PD are not relevant. Of course, Rossion knows these trials are needed, otherwise subjects would only have to respond “same” on every trial. This illustrates something critical: while Rossion is interested in an illusion he only perceives on same trials, his measurement of his participants' perception nonetheless depends on their decisions: whether to respond same or different on any given trial. Participants choose between two responses on each trial and their overall willingness to say “same” regardless of the trial type can vary. Rossion describes an ideal situation where an observer responds only on the basis of the illusion. According to him they would have no reason to respond “same” on trials where the relevant parts are different and the irrelevant parts are the same, apart from pressing the wrong button by mistake. At the same time, when they respond “different” to a trial where the relevant part is same and the irrelevant part is different, Rossion always interprets this as being due to holistic processing.

There is no question that such mistakes are possible, and that holistic processing can occur on same trials. But to argue that these are the *only* forces “yes” on tone present trials because they have really heard the tone, and that they can only respond “yes” on tone absent trials because their finger twitched.

Critically, in both a tone detection task and the composite paradigm, if errors are to be used, performance needs to be kept from ceiling. This can be accomplished in many ways, for instance by using stimuli that are relatively confusable or briefly presented. What this means is that sometimes a participant just does not really have a good sense for what the answer is, *but still must press a key*. Can Rossion guarantee that his participants will respond “same” exactly 50% of the time on these trials where they are guessing? Let's say they do, and Rossion has included a larger proportion of same trials without telling them, as he advocates. This alone would increase the number of same trials for which participants respond “different”, and would be interpreted as more evidence for holistic processing (see

Richler, Cheung & Gauthier, 2011a, Experiment 2). Or let's say they cannot respond "same" exactly 50% of the time when they are not sure, that some participants are more conservative than others: they would make fewer "same" responses than other subjects and that would be interpreted as less holistic processing. Or let's say that because of their experience with the other, easier, trials, participants get a sense that they need to be more conservative on aligned trials (where the irrelevant part is more difficult to ignore) than on misaligned trials: they may develop different responses biases for those aligned and misaligned trials on which they are more likely to guess. We could continue, but hopefully the point is sufficiently clear: it is one thing to predict what an observer will do when they have a clear percept, but a lot harder to predict what they will do when the task is sufficiently difficult.

Perhaps the clearest demonstration of this problem comes from manipulations where we influenced response bias through deception, and measured the impact on PD and CD measures of holistic processing. In Richler, Cheung & Gauthier (2011a) we tested whether response biases related to strategy would also influence holistic processing in the PD. We lied to participants about the proportion of same and different trials in the experiment, such that the 75%-same group were told that the correct response would be "same" on 75% of the trials, and the 75%-different group was told that the correct response would be "different" on 75% of the trials. Importantly, regardless of instructions, the true proportion of same and different trials in the experiment was always 50–50. As expected, the instructions influenced participants' response bias (see Figure 4, left panel): participants in the 75%-same group were more likely to respond "same", and participants in the 75%-different group were more likely to respond "different".

More importantly, false instructions also influenced whether evidence for holistic processing was observed in the PD (Figure 4, right panel): participants who were simply told that there were more "same" trials showed no alignment effect, and those who were told that more trials were "different" showed the reverse effect (better performance on aligned vs. misaligned trials)! It is difficult to explain how mere lies about the proportion of trials could eliminate the powerful illusion that Rossion describes. In fact, it does not, as it remains stable in the CD (a robust congruency × alignment, $\eta^2_p = .38$, with no three-way interaction with instructions, $\eta^2_p = .001$). But the measure Rossion champions, the PD, is clearly affected.

On the confound between congruency and response in the PD: "*there are actually no congruent trials!*"

Rossion also argues that congruency is not relevant to the PD because it was never manipulated in the first place. But while it was not *intentionally* manipulated, same trials in the PD are all incongruent and different PD trials are all congruent. Rossion seems to forget that participants can be influenced by any aspect of the design, even if he does not care about them. Note that in the CD, participants are no more instructed to attend to congruency than in the PD. If congruency matters in one design, there is no reason that it could not in the other. Put another way, if a research assistant had made a mistake and all the same trials happened to be shown in a smaller size than the different trials, even if this was not a dimension of interest for Rossion, he would have to acknowledge it as a confound.

Perhaps most strikingly, Rossion argues that there can be no influence of response biases if one *only* considers the same trials. It would be very convenient indeed if limiting our attention to hit rates eliminated response biases, as it would obviate the need for more complicated frameworks like signal detection theory. We encourage the reader to think twice before following such advice.

Now, why would the congruency status of these trials matter? There are well known conflict-monitoring systems in the brain that detect conflict, both at the response *and* the pre-response levels and increase cognitive control accordingly (Egner et al., 2007; Botvinick et al., 2001). Consequently, responses on same-incongruent trials can reflect a mix of holistic processing and of control processes engaged by the conflict elicited by the stimulus. The degree of conflict that is detected could differ as a function of alignment too, and as such we may observe different combinations of two distinct influences (one driven by holistic processing and one driven by control mechanisms) in each cell of the design. Even worse, it seems plausible that individual differences in cognitive control and holistic processing are independent (explaining why the PD measure performs particularly poorly in individual differences research – DeGutis et al., 2013; Konar et al., 2010; Richler, Cheung & Gauthier, 2011b) and that they may be modulated by completely different factors. In the PD, any contamination by response biases driven by congruency, or by its interaction with alignment, simply cannot be dissociated from holistic effects: neither using signal detection analyses on the PD or rhetorical arguments can change the fact that congruency and responses are completely confounded. The CD can separate failures of selective attention to the irrelevant part in aligned and misaligned conditions from other response biases. The researcher who fails to include the complementary trials required to separate performance from bias (i.e., the CD trials) on Rossion's advice would be well advised to consider whether Rossion's intuitions cover all possible influences during this task.

On the observation that the PD is susceptible to the influence of response bias: “This response bias is exactly what experimenters aim for”

Rossion argues at length that the response bias in the PD reflects holistic processing. He does seem to miss that the response bias we have argued to be problematic (e.g., Cheung et al., 2008; Richler, Cheung & Gauthier, 2011a,b; Richler, Mack et al., 2011) is the response bias that differs as a function of congruency and alignment, a bias that cannot be measured in the PD. Based on other comments Rossion makes, we could assume he feels the same about this response bias, since he argues nothing interesting happens in any of the cells of the CD but the same-incongruent trials.

Wondering if we missed something here, we decided to go back and re-analyze data from a face composite experiment that used the CD, and in which we found evidence of holistic processing using the congruency × alignment interaction in the CD ($F_{1,40} = 5.11$, $p < .05$), but no alignment effect in the PD ($t_{40} = 1.40$, $p = .17$). Perhaps some of our participants were confused due to the inclusion of all these experimental conditions Rossion deems useless? Perhaps this response bias which we argue should be set aside when quantifying holistic processing can in fact point us in the right direction? To determine whether finding evidence for holistic processing is dependent on response biases, we calculated the congruency × alignment interaction in response bias. Then, we used a median split to divide participants into two groups based on the magnitude of this interaction value, and calculated the PD alignment effect for each group.

Lo and behold, as predicted by Rossion, the interaction between the alignment effect and bias group is significant ($F_{1,39} = 17.25$, $p < .001$, $\eta^2_p = .31$; See Figure 5, right panel)! There is no alignment effect for Bias Group 1 ($t_{19} = 1.80$, $p = .09$), but a significant alignment effect – which would be interpreted as evidence for holistic processing in the PD – for Bias Group 2 ($t_{20} = 4.10$, $p = .001$). In contrast, holistic processing as indexed in the CD is significant and does not differ between groups ($F_{1,39} = .21$, $p = .65$, $\eta^2_p = .005$). This is not surprising since the congruency × alignment interaction was quantified in d', which is in principle independent from response bias.

We plotted response bias for both groups in the left panel of Figure 5 to understand exactly how biases differed between the two groups. To recap, Rossion argues that response bias is not problematic for the partial design because a bias to respond "different" on aligned trials *is* the illusion. As can be appreciated from the figure, in Bias Group 1 there is an effect of congruency on misaligned trials; participants are more likely to respond "different" on misaligned-incongruent trials. In contrast, in Bias Group 2 there is a larger effect of congruency on aligned trials – subjects in that group are more likely to respond "different" on aligned-incongruent trials.

Thus, as Rossion expected, a bias to respond "different" on aligned-incongruent trials – what Rossion argues "is" the composite illusion – allowed us to find participants who show consistent evidence of holistic processing… *with inverted faces* (this being the data from Experiment 1 in Richler, Mack et al., 2011). As discussed briefly in the last section below, Rossion is adamant that inverted faces should not be processed holistically. But, based on our data above, Rossion either has to concede that inverted faces can sometimes be processed holistically (a conclusion we would not dispute, although we draw it for very different reasons, namely the fact that we get a congruency × alignment effect in this experiment when we analyze the data for ALL subjects, not just half). Or, he must acknowledge that, at least sometimes, the response bias isn't what he's after.

On inversion: [holistic processing of inverted faces in the CD is] "incompatible with the disappearance of the composite illusion with inversion"

Rossion returns to inversion repeatedly throughout his paper, making too many claims for us to address them all here. For the reader interested in choosing a measure of holistic processing, what is critical is the claim that the PD is a better measure of holistic processing because it shows no holistic processing of inverted faces, which is consistent with his perceptual experience. The critical problem with this logic is that the choice of measurement should not be guided by whether or not an expected result is obtained. As psychologists we make and test predictions to further our understanding of the underlying processes we are interested in; if we already know how the results "should" turn out, what is the point of running experiments? Admittedly, sometimes we have a good basis on which to make predictions, and Rossion correctly cites a large literature that finds important differences between upright and inverted faces. Nobody argues that inversion has no effect on face perception. However, several have concluded that the difference is of a quantitative, not qualitative, nature (Loftus et al., 2004; Riesenhuber et al., 2004; Sekuler et al., 2004; Valentine & Bruce, 1988). Importantly, in Richler, Mack et al. (2011) we did find differences in the temporal dynamics of the effects for upright and inverted faces, even though both were ultimately processed holistically. There may still be ways in which the processing of upright and inverted faces differ at the longest presentation times even when they did not in the CD, and this would be a perfectly reasonable empirical question to ask. More importantly, we believe it is important to make an effort to separate theoretical considerations from those that motivate our decisions about measurement, for fear of being caught in a vicious circle.

CONCLUSION

To date, all the empirical studies that compare the PD and CD have concluded in favor of the CD (DeGutis et al., 2013; Cheung et al., 2008; Richler, Cheung & Gauthier, 2011,a,b; Richler, Mack et al., 2011). Nonetheless, like many disagreements about measurement, the debate between the PD and CD versions of the composite task has mostly played out outside

of public scrutiny in the depths of the peer-review process. We thank Rossion for bringing to light his reasons for continuing to use the PD despite the limitations we have been warning against, and we hope the reader interested in holistic processing will find this exchange useful in guiding their choice of measurement.

Acknowledgments

G and R would like to thank all their Cs, members of the OPL and CatLab, Matt Crump, Mike Mack, and Ben Tamber-Rosenau for providing comments on earlier versions of this manuscript, and Jackie Floyd, Emily Sauder and Bikang Zhang for assistance with data collection in the goat and bunny experiment. This work was supported by the NSF (Grant SBE-0542013), VVRC (Grant P30-EY008126) and NEI (Grant R01 EY013441-06A2).

Appendix

Appendix

(Abbreviations used: HP: holistic processing; CD: complete design; PD: partial design)

Claim	Reply
#2: *"Holistic face processing has received several definitions that are not fundamentally different from each other."*	This is an opinion. See Richler et al. (2012) for a different opinion about various operational definitions of HP.
#16: *"Part-based analytic processing is relatively well-preserved for inverted faces, but…perception of the individual face as a whole is impaired by inversion."*	See Richler, Mack et al. (2011) for data indicating that inverted faces are processed holistically, and McKone & Yovel (2009) for a review of feature inversion effects (i.e., impairment in recognition of individual features in inverted faces).
#22: *"Individuation of a nonface object from another member of the same category appears to rely essentially on part-based analysis."*	The papers cited in support of this claim do not take perceptual expertise into account.
#23: *"…studies that have applied the composite paradigm with nonface objects have failed to report any composite effects."*	See Bukach et al. (2010) and Wong et al.(2009a) for examples where HP was observed for non-face objects of expertise in the composite task.
#28: Holistic processing is supported by LSF (vs. HSF) information.	See Cheung et al. (2008) for evidence that HP is comparable for LSF and HSF faces in the complete design. Goffaux (2009) finds different result in a CD version that has no misaligned condition and uses simultaneous matching. Jury still out?
#38: *"A recent study (Wang, Li, Fang, Tian & Liu, 2012) …isolated the face-specificity measure of face recognition performance by subtracting performance at recognizing nonface objects. The composite face effect in correct RTs correlated significantly with this face-specific measure."*	In Wang et al. the individual correlations with the PD measure of HP 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 a relationship between the PD measure of HP and object processing, not face processing (see Humphreys, 1990, for a discussion of common errors in the interpretation of correlations with difference scores).
#41: *"Overall, the use of the composite face illusion in fMRI indicates that faces are represented holistically in face-sensitive areas of visual cortex"".*	See Gauthier & Tarr (2002) and Wong et al. (2009b) for evidence based on the CD that nonface objects of expertise are also represented holistically in the same brain areas.
#42: *"This early [N170] effect…identified the functional locus of the composite face effect at the earliest face perception stage, suggesting that facial-parts are not independently processed as face-like entities before being*	This interpretation of the N170 effect cannot be directly drawn from the data. Moreover, on the timescale of neurons, 170ms is not "fast" or "early" at all: neural responses in low-level visual areas begin as rapidly as 20ms post-stimulus onset (Casagrande et al., 2005), with feedback between higher and lower visual areas occurring in less than 50ms

Claim	Reply
integrated into a holistic representation."	(Pascal-Leone & Walsh, 2001).
#60: *"…since participants are not aware [that there are fewer different trials], it can only increase their tendency to respond `different' for `same' trials, leading to a higher proportion of mistakes and the chance to observe more clearly the composite face effect."*	This is an opinion about how to get a particular behavioral effect, not how to investigate a perceptual phenomenon.
#61: *"..if one includes many `different' trials in the study, participants might consider that in comparison to these real `different' trials, the illusory different top halves of faces do not look that different after all."*	This is an opinion about how to get a particular behavioral effect, not how to investigate a perceptual phenomenon.
#91: *"…behavioral studies using these other paradigms have provided information about holistic/configural face perception that generally agrees with studies using the composite face paradigm."*	See Richler et al. (2012) for a discussion of why different measures of HP based on different definitions need not produce the same results. See Wong et al. (2010) for data showing that one of the paradigms referred to here (Thatcher illusion) is actually not face-specific.
#115: *"…in the congruency/interference paradigm one cannot determine whether it is the incongruent context face half that interferes with the processing of the target face half, or if it is the congruent condition context face half that facilitates processing of the target face half*".	Interference vs. facilitation can be assessed by including isolated parts trials in the experiment (e.g., Richler, Bukach & Gauthier, 2009; Richler, Tanaka et al. 2008). See Richler, Tanaka et al. (2008) for evidence that the incongruent face half interferes with processing of the target part.
#119: *"…the results of [Richler et al., 2009b] do not fully support the authors' argument (Figure 1 of that paper) and suggest a contribution of a naming response conflict."*	This claim is based only on the accuracy data, which was at ceiling. Therefore, in that paper we placed greater emphasis on the RT data (see Richler, Cheung et al., 2009).
#121: *"[Richler & Gauthier] have been very critical of the misalignment manipulations in many publications, claiming that `the congruency effect provides a single measure of holistic processing without necessitating a misalignment manipulation to measure it' (Cheung et al., 2008, p. 1328)."*	A misaligned condition is *always* necessary in the PD, and in the CD it is not necessary in the same way, as a congruency effect can be measured and interpreted for aligned stimuli. We have in fact discussed very explicitly situations where the misaligned condition in the CD is particularly important (Richler, Wong & Gauthier, 2011). This critique is also out of touch with our published work: a misalignment manipulation was included in all but 4 (Richler, Gauthier et al., 2008, Experiment 1; Richler, Tanaka et al., 2008, Experiment 2; Richler, Bukach & Gauthier, 2009, Experiment 3; Richler, Mack et al., 2009) of the 16 experiments (in nine empirical papers) using the composite task that we (G&R) have published together.
#127: *"…because misaligned trials give rise to significant congruency effects at times of comparable magnitude to those observed for aligned trials, the authors claimed that misaligned faces were processed holistically (Richler et al., 2008b, experiment #1)."*	This is a misrepresentation of our results and conclusions. We found that misalignment at study had no effect on HP. Misalignment at test led to significant congruency effects that were significantly smaller in magnitude than those obtained for aligned faces. Therefore, we concluded that misalignment does not eliminate, but significantly attenuates, HP.
#128: *"Moreover, despite the fact that congruency effects for nonface objects (`Greebles') were equally large for aligned and misaligned stimuli (Gauthier et al., 1998; Gauthier & Tarr, 2002), these authors concluded that these nonface objects were processed holistically."*	Note that these experiments predate any formal comparison of the CD and PD. This critique focuses on a p=.05 level as a cliff, to suggest that effects were not existent in studies that were low-powered (n=10 in each case) because they were costly training studies. The focus is on effects associated with p-values of .06 and .09. The reader is invited to read the original papers, consider that effect sizes in short-term training studies are going to be smaller than for the same effects in real-world experts (as in car experts in Bukach et al., 2010). Much has been written about the perils of dichotomous thinking about

Claim	Reply
	significance, an interested reader could start with (Rosnow & Rosenthal, 1989).
#131: *"GRC do not only include both the `same' and `different' trials: the consider an interaction between two factors…but also a main congruency effect…and an alignment effect…as evidence for holistic processing." "The authors typically interpret all three kinds of effects instead of one".*	This is inaccurate. We have never used all three of these effects as possible places were we could find HP. Indeed, in all of our recent work (all papers published after 2009) we only measure HP using the congruency × alignment interaction (see Richler, Wong & Gauthier, 2011, for a review).
#137: Only misaligning the test face and not the study face is a methodological confound.	This is inaccurate. We only misalign faces at test because misalignment at study has been shown to produce contextual congruency effects in objects in novices (Richler, Bukach & Gauthier, 2009), but does not impact performance with faces (Richler, Tanaka et al., 2008). It is therefore an empirically motivated guard against the kind of spurious HP effects Rossion himself warns against (does he, or does he not, want us to be conservative about what we call HP?). Moreover, when we compare HP for faces to that for other object categories, or novices to experts, we use the same procedure, so this cannot account for differences we observe.
#140: *"However, with the exception of one study (Cheung et al., 2008), GRC also introduced another significant modification in their congruency/interference paradigm. That is, participants have to consider both halves of the study face".*	This is inaccurate. Participants also only made judgments on top face halves in Richler, Cheung & Gauthier (2011b) and McGugin et al. (2012).
#142: If the cued part is unknown at study *"the participant will certainly make more mistakes when there is an incongruent part anywhere in the second visual display presented (i.e., incongruent trials)…"*	This is an opinion. See Richler, Tanaka et al. (2008, Experiment 2) for data showing no effect of congruency when the target and task-irrelevant face halves are presented side-by-side at test. In other words, participants do not make more mistakes when there is an incongruent part anywhere in the test display. Moreover, when we compare HP for faces to that for other object categories, or novices to experts, we use the same procedure, so this cannot account for differences we observe.
#143: *"In the standard composite face paradigm the participant knows in advance which of the two halves…has to be encoded, and can therefore fixate gaze accordingly…If presentation time is longer, [the participant] will probably alternate between fixating one of the two halves of the study face…thus, you cannot fairly compare the effects observed when presentation duration is variable at encoding, especially when you compare conditions allowing only one fixation (e.g., 17ms to 183ms) to conditions allowing several fixations (>183ms until 800ms; Richler et al., 2009a)."*	This is precisely the point of the manipulation of presentation time: to compare the magnitude of HP when there is more time available to potentially encode both object parts even though only one is relevant. Importantly, sensitivity (d') does not differ between conditions when presentation time of the study face is limited and conditions when presentation time of the test face is limited (Richler, Mack et al., 2009).
#146: The complete design has too many trials. *"This is not parsimonious at all, and violates a general principle in research that one should include only the conditions in the paradigm that allow to test specific hypotheses rather than manipulate all possible variables and then expect some regularizes (`laws') to emerge."*	We resent the implication here that we manipulate "everything" and then "data mine" to find significant effects. We believe that all the trials in the CD are necessary if we want to test specific hypotheses about HP and guard against bias confounds.
#147: *"Such a high number of conditions is particularly problematic when one needs to assess holistic face processing in (very) young children…or in brain-damaged patients."*	This is an opinion. In the CD, 160 trials are typically necessary to assess HP (e.g., Richler, Cheung & Gauthier, 2011a,b). In de Heering, Houthuys & Rossion (2007) children are tested on 100 trials, and in Ramon, Busigny & Rossion (2010) patients are tested on 116 trials. This

Claim	Reply
	hardly seems like a dramatic increase, and frankly, it seems hardly ethical to test anyone at all in measures that have demonstrated low validity.
#150: *"…weaker correlation [between holistic processing and the CFMT] reported by DeGutis et al. (2013)."*	This is misleading. The correlation we reported in Richler, Cheung & Gauthier (2011b) was .40; the correlation reported by Degutis et al. was .33. Moreover, this result has also been replicated in McGugin et al. (2012), where the first-order correlation was .26, and the partial correlation (factoring out age, sex, and their interactions with holistic processing) was .33.
#152: *"Researchers in this field are not interested in the general processes that can drive such inter-individual correlations, but rather in what specifically differs between upright face processing, namely holistic face perception."*	Rossion seems to be confusing his own research interests with those of an entire field. A very cursory review found 16 articles published from 2010 onwards that use an individual differences approach in the study of face recognition: Avidan et al. (2011), Bukach et al. (2012), Davis et al. (2011), DeGutis et al. (2013), Dennett et al. (2011), Dennett et al. (2012), Germine et al. (2011), Konar et al. (2010), McGugin et al. (2012), Mondloch & Desjarlais (2010), Richler et al. (2011), Wang et al. (2012), Wilhelm et al. (2010), Wilmer et al. (2010), Zhou et al. (2012), Zhu et al. (2010).
#154: *"…the study of Mack et al. (2011) entitled `Indecision on decisional separability' dismissed entirely the conclusions reached by Richler et al. (2008a) that holistic processing has a decisional locus."*	In Mack et al. (2011) we explained why the analysis tool developed by Kadlec & Townsend (1992) and used in Richler, Gauthier et al. (2008) is not valid. Indeed, we only used that analysis in that single paper because we discovered these issues and made them public immediately.
#159: *"…Richler et al. (2009a) did not use any misaligned faces or inverted faces as a control, which makes it impossible to interpret their effects."*	This is an opinion, and the term "impossible" seems overly strong. Note also that we ourselves acknowledged this shortcoming, and addressed it in Richler, Mack et al. (2011, Experiment 2).
#163: *"Inverted faces are usually presented for reasonably long durations, so that Richler et al's (2011) claim that inverted faces would not be processed holistically only at short durations is clearly inconsistent with that literature."*	This is misleading. Studies that find no HP for inverted faces presented for long durations all use the PD (or simultaneous presentation, Goffaux, 2009). It is therefore not surprising that our results using the CD are incompatible because these two designs often yield different results.
#164: *"GRC's congruency/interference paradigm was initially developed to study the processing of nonface objects…and was applied only later to faces…"*	This is inaccurate. The CD was first used by Farah et al. (1998) with faces.
#165: The conclusion that objects of expertise are processed holistically are only based on congruency effects (Gauthier et al, 1998; 2002; 2003).	This is inaccurate. Alignment effects are reported in Gauthier & Tarr, (2002).
#166: *"[In Wong et al. 2009]…On RTs, there was a non-significant trend of an interaction between expertise, congruency and alignment…*	This is rhetorical. This three-way interaction is at p=.053. See answer to claim 128. We hope that enough expertise training studies will eventually have been done that a meta-analytical approach to thinking about what is going on is possible.
#167: *However, this advantage for the congruent-aligned condition was larger in novices than experts on sensitivity, pointing to a speed-accuracy trade-off in the task (Figure 5 of Wong et al., 2009)…"*	Rossion is comparing the Individuation training group to the No training group and apparently suggesting that a group × congruency × alignment interaction in d' that is non-existent (F=.0004) can explain an effect in RTs (F=3.33).
#168: *"This is problematic because if one aims at demonstrating that faces do not call upon specific holistic processes, the proper approach*	We never aimed at demonstrating faces do not elicit special HP. Many of our studies aim at testing whether HP increases with expertise.

Claim	Reply
is to test visual experts with nonface object categories by means of the very same paradigm used to obtain the strongest face-specific effects in novices."	We believe the proper approach is not to use the design that gives the "strongest" effect with faces, but the "cleanest" effect, one that can be validly associated with HP rather than response biases.

References

Alsius A, Navarra J, Soto-Faraco S. Attention to touch weakens audiovisual speech integration. Experimental Brain Research. 2007; 183(3):399–404.

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

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. Psychological Review. 2001; 108:624–652. [PubMed: 11488380]

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.

Bukach C, Cottle J, Ubiwa J, Miller J. Individuation Experience Predicts Other-Race Effects in Holistic Processing. Journal of Vision. 2012; 12(9):637–637.

Casagrande VA, Sary G, Royal D, Ruiz O. On the impact of attention and motor planning on the lateral geniculate nucleus. Progress in Brain Research. 2005; 149:11–29. [PubMed: 16226573]

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

Davis JM, McKone E, Dennett H, O'Connor KB, O'Kearney R, Palermo R. Individual differences in the ability to recognise facial identity are associated with social anxiety. PLoS ONE. 2011; 6(12):e28800. doi:10.1371/journal.pone.0028800. [PubMed: 22194916]

De Heering A, Houthuys S, Rossion B. Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. Journal of Experimental Child Psychology. 2007; 96:57–70. [PubMed: 17007869]

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. [PubMed: 23084178]

Dennett HW, McKone E, Edwards M, Susilo T. Face Aftereffects Predict Individual Differences in Face Recognition Ability. Psychological Science. 2012; 23(11):1279–1287. doi: 10.1177/0956797612446350. [PubMed: 23073026]

Dennett HW, McKone E, Tavashmi R, Hall A, Pidcock M, Edwards M, Duchaine B. The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. Behavior Research Methods. 2011; 44(2):587–605. doi:10.3758/s13428-011-0160-2. [PubMed: 22012343]

Egner T, Delano M, Hirsch J. Separate conflict-specific cognitive control mechanisms in the human brain. Neuroimage. 2007; 35:940–948. [PubMed: 17276088]

Farah MJ, Wilson KD, Drain M, Tanaka JW. What is "special" about face perception? Psychological Review. 1998; 105:482–498. [PubMed: 9697428]

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

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. [PubMed: 11999864]

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. [PubMed: 9798007]

Germine LT, Duchaine B, Nakayama K. Where cognitive development and aging meet: face learning ability peaks after age 30. Cognition. 2011; 118(2):201–210. doi:10.1016/j.cognition.2010.11.002. [PubMed: 21130422]

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. [PubMed: 16846295]

Humphreys LG. Erroneous interpretation of difference scores: Application to a recent example. Intelligence. 1990; 14:231–233.

Kadlec, H.; Townsend, JT. Signal detection analysis of dimensional interactions. In: Ashby, FG., editor. Multidimensional models of perception and cognition. Lawrence Erlbaum Associates; Hillsdale: 1992. p. 181-228.

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

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

Mack ML, Richler JJ, Gauthier I, Palmeri TJ. Indecision on decisional separability. Psychonomic Bulletin & Review. 2011; 18(1):1–9. [PubMed: 21327365]

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. [PubMed: 19815781]

McGugin RW, Richler JJ, Herzmann G, Speegle M, Gauthier I. The Vanderbilt Expertise Test and the contribution of general object recognition abilities to face recognition. Vision Research. 2012; 69:10–22. [PubMed: 22877929]

McManus JN, Li W, Gilbert CD. Adaptive shape processing in primary visual cortex. Proceedings of the National Academy of Sciences. 2011; 108(24):9739–9746.

Mondloch CJ, Desjarlais M. The function and specificity of sensitivity to cues to facial identity: An individual-differences approach. Perception. 2010; 39:819–829. [PubMed: 20698476]

Pascual-Leone A, Walsh V. Fast backprojections from the motion to the primary visual area necessary for visual awareness. Science. 2001; 292(5516):510–512. [PubMed: 11313497]

Ramon M, Busigny T, Rossion B. Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. Neuropsychologia. 2010; 48:933–944. [PubMed: 19944710]

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

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.

Richler JJ, Cheung OS, Gauthier I. Holistic processing predicts face recognition. Psychological Science. 2011b; 22:464–471. [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. [PubMed: 19293091]

Richler, JJ.; Gauthier, I. A meta-analysis of holistic face processing. (manuscript in preparation)

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.

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

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

Richler JJ, Palmeri TJ, Gauthier I. Meanings, mechanisms, and measures of holistic processing. Frontiers in Psychology. 2012; 3

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.

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. [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.

Rosnow RL, Rosenthal R. Statistical procedures and the justification of knowledge in psychological science. American Psychologist. 1989; 44(10):1276.

Rossion B. The composite face illusion: A whole window into our understanding of holistic face perception. Visual Cognition. 2013

Rossion B, Pourtois G. Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. Perception. 2004; 33(2):217–236. [PubMed: 15109163]

Sekuler AB, Gaspar CM, Gold JM, Bennett PJ. Inversion leads to quantitative, not qualitative, changes in face processing. Current Biology. 2004; 14:391–396. [PubMed: 15028214]

Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of Cognitive Neuroscience. 2005; 17(7):1098–1114. [PubMed: 16102239]

Valentine T, Bruce V. Mental rotation of faces. Memory and Cognition. 1988; 16:556–566. [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. [PubMed: 22222218]

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. [PubMed: 20677889]

Wilmer JB, Germine L, Chabris CF, Chatterjee G, Williams M, Loken E, Nakayama K, Duchaine B. Human face recognition ability is highly heritable. Proceedings of the National Academy of Sciences. 2010; 107:5238–5241.

Wong AC-N, Palmeri TJ, Gauthier I. Conditions for face-like expertise with objects: Becoming a Ziggerin expert – but which type? Psychological Science. 2009a; 20:1108–1117. [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. 2009b; 4:e8405. [PubMed: 20027229]

Wong YK, Twedt E, Sheinberg D, Gauthier I. Does Thompson's Thatcher effect reflect a face-specific mechanism? Perception. 2010; 39:1125–1141. [PubMed: 20942363]

Zaretskaya N, Anstis S, Bartels A. Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. The Journal ofNeuroscience. 2013; 33(2):523–531.

Zhou G, Cheng Z, Zhang X, Wong AC-N. Smaller holistic processing of faces associated with face drawing experience. Psychonomic Bulletin & Review. 2012; 192:157–162. [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. [PubMed: 20036540]

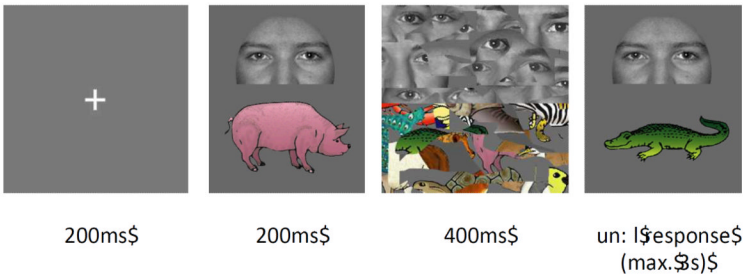

Figure 1.

Twenty-two participants from the Vanderbilt University community received monetary compensation for participation. Data from two participants were discarded for the use of incorrect response keys. Data from a third participant were discarded for mean RTs more than 2 SD below the mean (mean RT < 250ms). Stimuli were 20 face tops (10 male, 10 female; Goffaux & Rossion, 2006) and 20 color line drawings of animals (Rossion & Pourtois, 2004). Note that in addition to the example animals shown here, we also used the goat and bunny images shown in Figure 37 of Rossion (2013). A sample trial sequence is illustrated above. Timing of elements was based on Figure 2 of Rossion (2013). On each trial, participants were instructed to judge whether the face half was the same (`J' key) or different (`K' key). There were 20 trials for each combination of congruency (congruent/incongruent), alignment (aligned/misaligned) and correct response (same/different) for a total of 160 trials. The experimental block was preceded by a 16-trial practice block. On "different" trials, the sex of the test face half was the same as the study face half. On misaligned trials, the study image was aligned, and the animal image was moved 128 pixels to the left in the test image, so that the target face half was presented in the same physical location in the study and test displays, as recommended by Rossion.

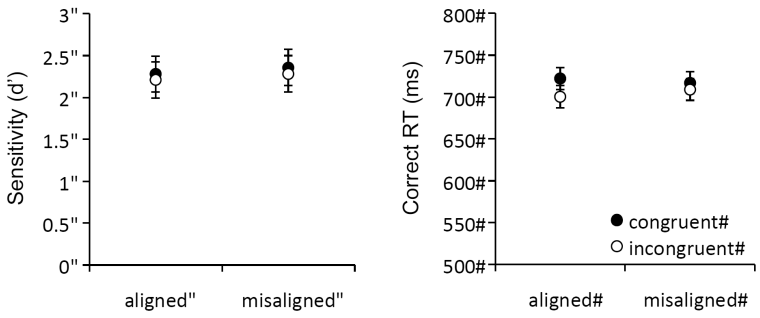

Figure 2.
Sensitivity (d'; left panel) and Correct RT (ms; right panel) as a function of congruency and alignment. Error bars show 95% confidence intervals of within-subjects effects.

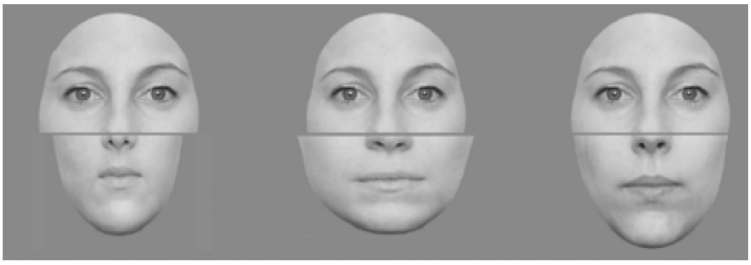

Figure 3.
The same face top is aligned with a bottom face half of different widths. The top and bottom half are only perfectly aligned in the face on the right.

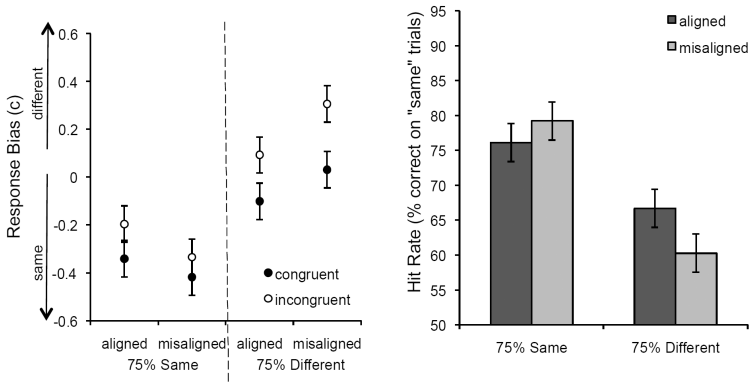

Figure 4.

Re-plotted data from Experiment 1 in Richler, Cheung & Gauthier (2011a). Left Panel: Response bias (c) on congruent and incongruent trials as a function of alignment for each instruction group. Positive values indicate a bias to respond “different” and negative values indicate a bias to respond “same”. Error bars show 95% confidence intervals of the interaction. Right Panel: Hit rate (percent correct on “same”-incongruent trials) as a function of alignment for each instruction group. Error bars show 95% confidence intervals of the interaction.

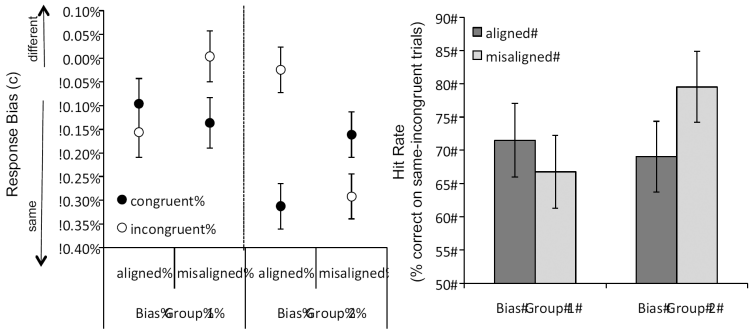

Figure 5.

Left Panel: Response bias (c) on congruent and incongruent trials as a function of alignment for each bias group. Positive values indicate a bias to respond “different” and negative values indicate a bias to respond “same”. Error bars show 95% confidence intervals of the within-subject effects. Right Panel: Hit rate (percent correct on “same”-incongruent trials) as a function of alignment for each bias group. Error bars show 95% confidence intervals of the within-subjects effect.