Research Report

Psychological Science
22(4) 464–471

DOI: 10.1177/0956797611401753
http://pss.sagepub.com

Holistic Processing Predicts Face Recognition

Jennifer J. Richler[1], Olivia S. Cheung[2], and Isabel Gauthier[1]
[1]Vanderbilt University and [2]Martinos Center, Massachusetts General Hospital, Harvard Medical School

Abstract

The concept of holistic processing is a cornerstone of face-recognition research. In the study reported here, we demonstrated that holistic processing predicts face-recognition abilities on the Cambridge Face Memory Test and on a perceptual face-identification task. Our findings validate a large body of work that relies on the assumption that holistic processing is related to face recognition. These findings also reconcile the study of face recognition with the perceptual-expertise work it inspired; such work links holistic processing of objects with people's ability to individuate them. Our results differ from those of a recent study showing no link between holistic processing and face recognition. This discrepancy can be attributed to the use in prior research of a popular but flawed measure of holistic processing. Our findings salvage the central role of holistic processing in face recognition and cast doubt on a subset of the face-perception literature that relies on a problematic measure of holistic processing.

Keywords

face perception, individual differences, holistic processing

Received 9/3/10; Revision accepted 12/10/10

Face recognition challenges perception because similar facial features are arranged in similar configurations on all human faces. As such, subtle differences in facial features and their spatial relations are particularly useful for discriminating faces. To facilitate extraction of configural information, people process faces holistically, as evidenced by the fact that it is more difficult to ignore part of a face than part of an object (Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Le Grand, & Mondloch, 2002; Young, Hellawell, & Hay, 1987). Accordingly, people's ability to discriminate and recognize faces should depend at least in part on holistic processing.

Surprisingly, holistic processing and face-recognition ability have never been linked empirically. Support for the relationship between holistic processing and face-recognition ability is mainly indirect, coming from studies in which perceptual experts with superior object-identification ability also demonstrate holistic processing in their domain of expertise (Bukach, Phillips, & Gauthier, 2010; Gauthier & Tarr, 2002; Wong, Palmeri, & Gauthier, 2009). However, in a recent article, Konar, Bennett, and Sekuler (2010) argued that holistic processing does not predict face-identification ability. It is important to examine this issue further because holistic processing plays a pivotal role in studies of face recognition. Studies have used holistic processing to track the development of face recognition (e.g., Macchi Cassia, Picozzi, Kuefner, Bricolo, & Turati, 2009; Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007), to study abnormal development of face recognition (e.g., among individuals with developmental prosopagnosia; Le Grand et al., 2006) or populations with face-recognition deficits that are part of more widespread cognitive impairment (e.g., schizophrenia; Schwartz, Marvel, Drapalski, Rosse, & Deutch, 2002), and to evaluate computational models of face recognition (Dailey & Cottrell, 1999). If holistic processing does not relate to performance recognizing faces, such efforts may constitute wild-goose chases.

Konar et al. (2010) suggested that their failure to find a relationship between face identification and holistic processing could be related to the specific nature of the tasks they used. We followed up on this concern by reassessing the relationship between holistic processing and face processing. In particular, we addressed two key issues.

First, we have questioned elsewhere the validity of the composite design used by Konar et al. (2010; this design was adapted from a naming task with familiar faces devised by Young et al., 1987; see also Hole, 1994) because of its

Corresponding Author:
Jennifer J. Richler, Department of Psychology, Vanderbilt University, 301 Wilson Hall, 111 21st Ave. South, Nashville, TN 37240
E-mail: jennifer.j.richler@vanderbilt.edu

susceptibility to response biases (e.g., Cheung, Richler, Palmeri, & Gauthier, 2008; Richler, Mack, Palmeri, & Gauthier, 2011). In the study reported here, we used a measure of the composite task that is arguably more valid than the design used by Konar et al. (2010) and that has been related to expertise for objects (A.C.-N. Wong, Palmeri, & Gauthier, 2009).

Second, in everyday face recognition, an encountered face must be compared with many representations stored in memory to determine identity. Measures of face processing in a task in which participants need only to match faces within each trial, as used by Konar et al. (2010), may overestimate the contribution of featural strategies that are less available in real-world situations. Therefore, a task in which multiple target faces are stored in long-term memory may tap into the robustness of stored face representations and better represent individual differences relevant to everyday face recognition. For example, in a recent study, Furl, Garrido, Dolan, Driver, and Duchaine (2010) found that although a face-memory task was associated more strongly with a face-processing factor (as determined by principal component analysis), a perceptual face-matching task was associated more strongly with an object-processing factor. To sample individual differences in face processing better than Konar et al. (2010) did, we used both the face-matching task used by Konar et al. and the Cambridge Face Memory Test (CFMT), a well-validated measure of individual differences in face recognition (Duchaine & Nakayama, 2006).

Method

Thirty-eight members of the Vanderbilt University community (11 male, 27 female) ranging in age from 18 to 40 years (median age = 20 years) were compensated for participation in the study. Participants completed three tasks in the following order: the CFMT, the composite task, and the face-identification task. The study was approved by the local institutional review board.

CFMT

At the start of the CFMT, participants studied frontal views of six target faces for a total of 20 s. Then they completed an 18-trial introductory learning phase, after which they were presented with 30 forced-choice test displays. Each display contained one target face and two distractor faces. Participants were told to select the face that matched one of the original six target faces. The matching faces varied from their original presentation in terms of lighting condition, pose, or both. Next, participants were again presented with the six target faces to study, followed by 24 test displays presented in Gaussian noise. All trials were combined for each participant to yield a single measure of accuracy.

Composite task

Stimuli in the composite task consisted of 20 female faces. These images were cropped to create 20 face tops and 20 face bottoms. Top and bottom halves were randomly combined on every trial to form composite faces 256 × 256 pixels in size (see Fig. 1b). A white line (3 pixels thick) separated the face halves, resulting in a stimulus 256 × 259 pixels in size. The white line ensured that it was completely unambiguous where the top face half ended and the bottom half began, and this, if anything, was expected to facilitate selective attention.

On each of 160 trials, participants were instructed to judge whether the top half of the test face was the same as or different from the top half of the study face while ignoring the other, irrelevant bottom half. On 80 trials, the top and bottom halves were aligned, and on the other 80 trials, the halves were misaligned. In misaligned trials, the top half of the test face was moved 35 pixels to the left, and the bottom half was moved 35 pixels to the right; thus, the edge of the top half always fell on the center of the bottom half (see Fig. 1b for examples of stimuli and trial sequences).

There were four trial types in the composite task (see Fig. 1a). Two types were *same* trials, in which the relevant halves of the study and test faces were the same. In the two types of *different* trials, the relevant halves of the faces were different. Within *same* trials and *different* trials, faces could also be congruent or incongruent. In congruent trials, the irrelevant half was associated with the same response as the relevant half. In incongruent trials, the irrelevant face half was associated with a different response than the relevant half. Therefore, in congruent *same* trials, the irrelevant half of the test face was the same as the irrelevant half of the study face. In incongruent *same* trials, the irrelevant halves of the test and study faces were different. In congruent *different* trials, the irrelevant half of the test face was different from the irrelevant half of the study face. In incongruent *different* trials, the irrelevant halves of the study and test face were the same.

Two versions of the sequential-matching composite task have been used in previous research: the partial design and the complete design (see Fig. 1a). The partial design, used by Konar et al. (2010), consists of only two types of trial: congruent *different* and incongruent *same*. In *same* trials, an alignment effect indexes holistic processing: Accuracy is greater or reaction time (RT) is faster in misaligned trials than in aligned trials (Macchi Cassia et al., 2009; de Heering, Houthuys, & Rossion, 2008; Goffaux & Rossion, 2006; Hole, 1994; Le Grand et al., 2006; McKone & Robbins, 2007; Michel, Rossion, Han, Chung, & Caldara, 2006; Mondloch et al., 2007).

The complete design includes the two partial-design trials plus congruent *same* and incongruent *different* trials. Participants' failure to selectively attend to parts of faces is indexed by a congruency effect: Performance is better in congruent trials than in incongruent trials (Cheung et al., 2008; Farah et al., 1998; Gauthier, Curran, Curby, & Collins, 2003; Goffaux, 2009; Richler, Mack, Gauthier, & Palmeri, 2009; Richler, Tanaka, Brown, & Gauthier, 2008). Misalignment reduces the congruency effect (Cheung et al., 2008; Richler et al., 2008), and this interaction between congruency and alignment is particularly sensitive to expertise-driven holistic processing

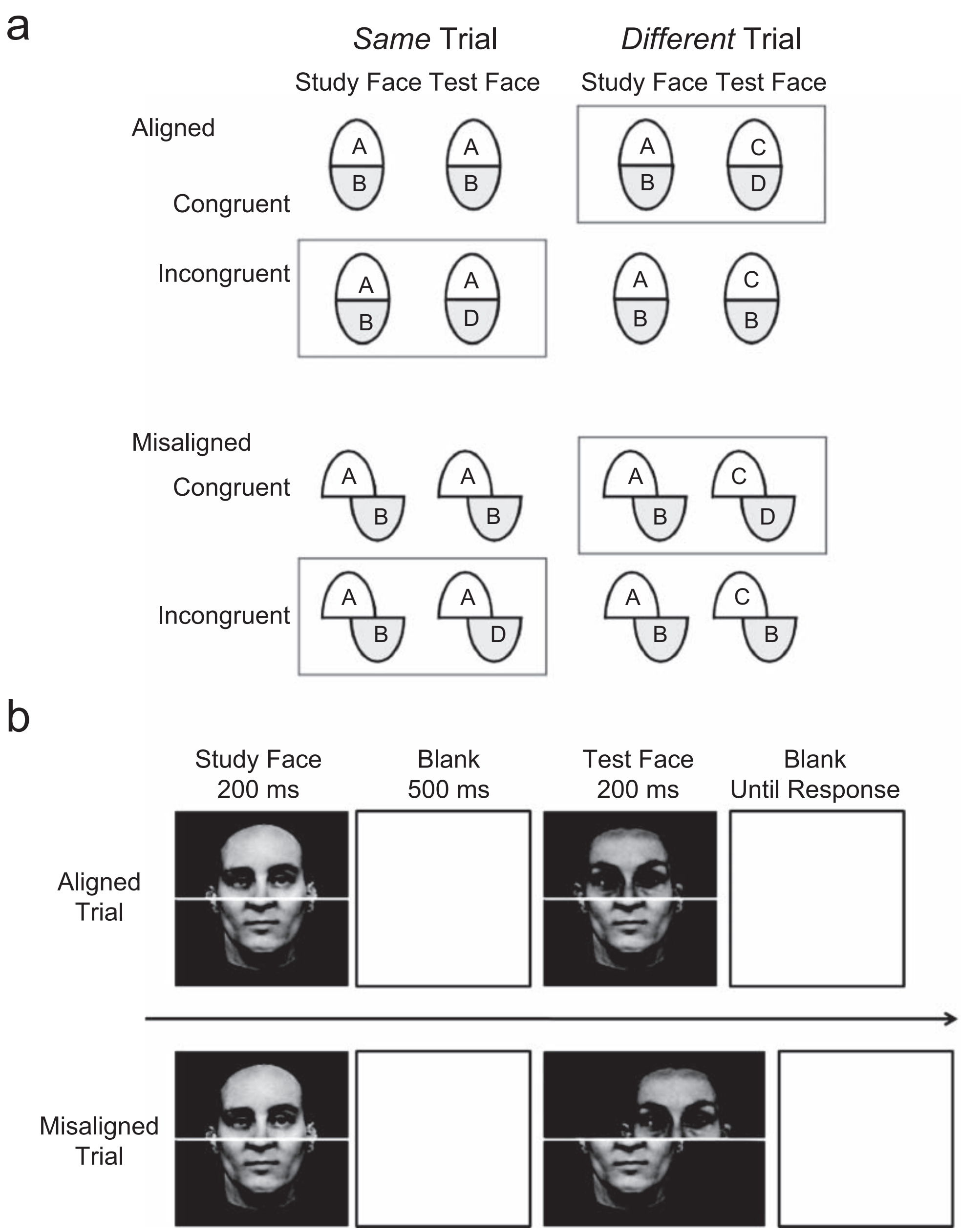

Fig. 1. Design of the composite task and sample trial structure. In the schematic diagram (a), letters represent facial identities. Task-relevant face halves are shown in white, and task-irrelevant halves are shown in gray. In *same* trials, task-relevant halves of the study and test faces were the same; in *different* trials, task-relevant halves were different. Both types of trials featured congruent and incongruent conditions. In congruent *same* trials, the irrelevant halves of the study and test faces were the same; in incongruent *same* trials, the irrelevant halves were different. In congruent *different* trials, the irrelevant halves of the study and test faces were different; in incongruent *different* trials, the irrelevant halves were the same. Face halves were presented aligned or misaligned. In the partial-design version of this task, only the trial types outlined in the gray boxes were presented; in the complete design, all trial types were presented. The examples in (b) illustrate the stimuli and trial sequence.

(Richler, Bukach, and Gauthier, 2009; A.C.-N. Wong, Palmeri, & Gauthier, 2009).

The partial-design measure was the first index of holistic processing used in the composite task (Hole, 1994; Young et al., 1987), but subsequently, both partial and complete designs have been extensively used (see the Supplemental Material available online). In the study reported here, we used the complete design, which gave us the flexibility to perform partial-design as well as complete-design analyses (see Cheung et al., 2008; Richler et al., 2011).

Face-identification task

Our face-identification task was modeled after the task used by Konar et al. (2010). On each of 120 trials, a target face was presented (200 ms). Participants then viewed a four-face display and had to select the face that matched the target face. This display was shown until a response was made. Target and foil faces were either all male (60 trials) or all female (60 trials) and differed in lighting conditions to prevent image matching.

Results and discussion

Holistic processing measured in the complete design of the composite task was observed in the group-level data, as revealed by a significant interaction between alignment and congruency in the analysis of d', $F(1, 37) = 5.28$, $p = .027$. This interaction was not significant in the analysis of RT, $F(1, 37) = 3.36$, $p = .075$.

Partial analyses revealed no significant effect of alignment in the group-level data in the analysis of accuracy, $t(37) = 1.31$, $p = .198$, or of RT, $t(37) = -1.587$, $p = .121$. The failure to find an alignment effect is not the result of running partial analyses on data collected in the complete design (see the Supplemental Material). Moreover, all other measures suggest that our participants were typical and processed faces holistically according to the complete design: The absence of an alignment effect in partial analyses may reflect the poor reliability of this measure of holistic processing.

Next, we examined correlations between holistic processing and measures of face-recognition ability. For each correlation, 95% confidence intervals (CIs) are reported. Significant correlations were still significant using Spearman correlations and after removing outliers (see the Supplemental Material).

Average accuracy on the CFMT and face-identification task was 76.35% (SD = 14.39%) and 75.13% (SD = 10.55%), respectively. Performance on these tasks was strongly but not perfectly correlated, $r(38) = .702$, CI = [.448, .815], $p < .0001$; this finding perhaps indicates an upper limit between each of these measures and holistic processing.

Using partial analyses within the complete design, we found that the magnitude of the alignment effect in RT did not correlate with face recognition—CFMT: $r(38) = .128$, CI = [−.199, .430], $p = .445$; face-identification task: $r(38) = .160$, CI = [−.168, .456], $p = .336$; nor did the alignment effect in accuracy correlate with face recognition—CFMT: $r(38) = .190$, CI = [−.138, .480], $p = .252$; face-identification task: $r(38) = .093$, CI = [−.233, .400], $p = .579$ (see Fig. 2). Furthermore, the alignment effect did not correlate with the alignment effect indexed using d', which was the measure used by Konar et al. (2010)—CFMT: $r(38) = -.074$, CI = [−.384, .251], $p = .660$; face-identification task: $r(38) = .040$, CI = [−.283, .335], $p = .809$. In sum, our partial-design analyses replicated Konar et al.'s (2010) findings: We found no evidence that holistic processing is linked to face recognition.

In contrast, holistic processing in the complete design predicted individual differences in face recognition (Fig. 3). Performance on the CFMT was significantly correlated with the magnitude of the Congruency × Alignment interaction in analyses of both d', $r(38) = .396$, CI = [.088, .635], $p = .014$, and RT, $r(38) = .334$, CI = [.017, .590], $p = .040$. Performance on the face-identification task was significantly correlated with holistic processing in analyses of RT, $r(38) = .482$, CI = [.192, .694], $p = .002$, but not of d', $r(38) = .031$, CI = [–.291, .347], $p = .851$. At least one prior study found that face matching and face memory differentially correlate with speed and accuracy in face recognition (Wilhelm et al., in press), but it is also possible that the RT measure of holistic processing is sometimes more sensitive than the d' measure (e.g., in A.C.-N. Wong, Palmeri, & Gauthier, 2009, only the RT index correlated with right fusiform gyrus activity). Holistic processing has been traditionally indexed by either or both of these dependent variables, and it is important to note that we found no trade-off between the two.

In addition to holistic processing, featural processing may contribute to face-recognition performance. Two multiple regression analyses were conducted, with performance on the face-identification task and on the CFMT as dependent variables. The four predictors in the model were the Congruency × Alignment interaction in the analysis of d' and of RT, and performance in the analysis of d' and of RT in the misaligned-faces conditions (averaging across congruency). Performance for misaligned trials provides an estimate of featural processing because when face parts are misaligned, selective attention to a part is more successful, as evidenced by the smaller congruency effect in misaligned trials. Consistent with our conjectures about the differences between the face-identification task and CFMT, we found independent contributions of the Congruency × Alignment interaction in the analysis of RT and of d' for misaligned trials on face-identification scores, but only the Congruency × Alignment interaction in d' was a significant predictor of CFMT scores (Table 1). Performance on both tasks relies on holistic processing, but the CFMT allows little or no contribution from featural processing.

Holistic processing measured with the complete design in the composite task predicted individual differences in face recognition: The larger the effect of holistic processing (Congruency × Alignment interaction), the better the face-recognition performance. This finding reconciles the idea that holistic processing is important to face processing with studies linking holistic processing and perceptual expertise (Gauthier & Tarr, 2002; A.C.-N. Wong, Palmeri, & Gauthier, 2009). In fact, because the face-identification task and the CFMT are similar to measures of expertise in nonface domains (Bukach et al., 2010; Gauthier et al., 2003; Y.K. Wong & Gauthier, 2010), our results suggest that holistic processing predicts expertise for both faces and nonface objects.

Why do the complete- and partial-design measures of holistic processing lead to different conclusions? One reason is that the partial-design measure does not take into account possible

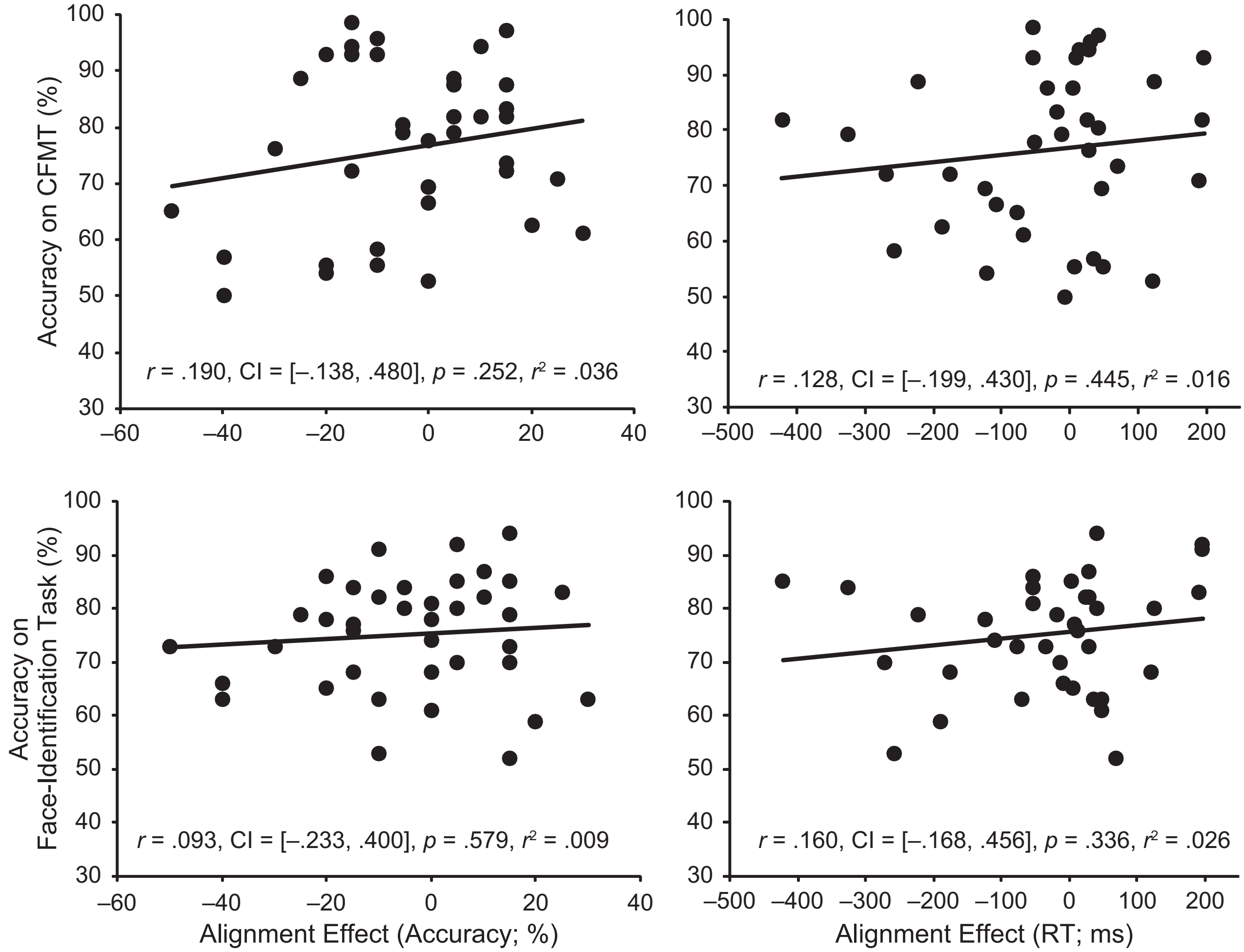

Fig. 2. Scatter plots (with best-fitting regression lines) showing correlations between measures of holistic processing in the partial design of the composite task (*x*-axes) and face-identification ability (*y*-axes). Results are shown for the Cambridge Face Memory Test (CFMT; top row) and the face-identification task (bottom row). Holistic processing in the partial design was indexed by the alignment effect (difference in performance between misaligned and aligned incongruent *same* trials) in accuracy (left column) and reaction time (RT; right column). CI = confidence interval.

influences of response biases, whereby participants choose to respond "same" more often in some conditions regardless of their ability to discriminate the face halves (Cheung et al., 2008; Richler, Mack, et al., 2009; Richler et al., 2011). Indeed, although d', as used by Konar et al. (2010), provides a discriminability measure that is independent of response bias, its use does not fully resolve this issue. In the partial design, irrelevant face halves are always different, therefore *same* trials are always incongruent and *different* trials are always congruent. In the complete design, congruency often influences response bias, and often differentially on aligned versus misaligned trials (Cheung et al., 2008; Richler, Mack et al., 2009; Richler et al., 2011). A limitation of the partial design is that the alignment effect is confounded with congruency, and there is no way to measure the bias associated with congruency and how this bias is modulated by alignment. In contrast, in the complete design, d' as a function of both alignment and congruency is robust to manipulations that influence response bias (Cheung et al., 2008; Richler, Mack, et al., 2009; Richler et al., 2011). Indeed, in the study reported here, response bias was correlated with the magnitude of the alignment effect in analyses of accuracy and d'—accuracy: $r(38) = -.523$, CI = [−.721, −.245], $p = .001$; d': $r(38) = -.666$, CI = [−.44, −.812], $p < .0001$—but not holistic processing measured in the complete design, $r(38) = -.280$, CI = [−.550, .043], $p = .088$.

General Discussion

The fact that faces are processed holistically is what makes them special—face perception relies on holistic processing more than object perception to maximize sensitivity to configural information (Farah et al., 1998). A considerable amount of research depends not only on the validity of the way in which holistic processing is measured, but even more fundamentally on the assumption that holistic processing is relevant to understanding face processing (e.g., Macchi Cassia et al., 2009; Dailey &

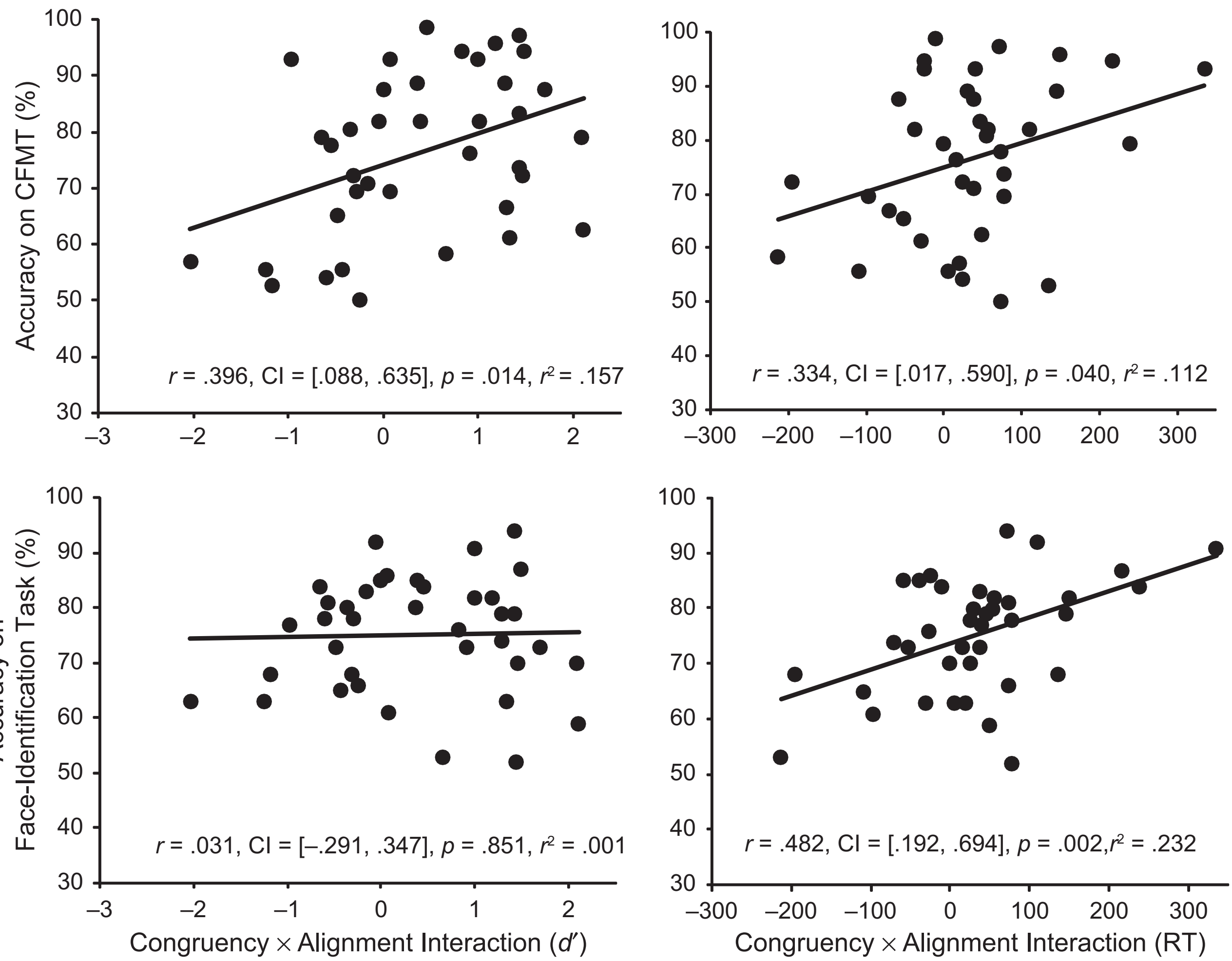

Fig. 3. Scatter plots (with best-fitting regression lines) showing correlations between measures of holistic processing in the complete design of the composite task (x-axes) and face-identification ability (y-axes). Results are shown for the Cambridge Face Memory Test (CFMT; top row) and the face-identification task (bottom row). Holistic processing in the complete design was indexed by the magnitude of the Congruency × Alignment interaction in d' (calculated as z scores for hits minus z scores for false alarms; left column) and reaction time (RT; right column). CI = confidence interval.

Table 1. Results of Multiple Regression Analyses

Model and predictor	β	*SE*	*t*	*p*
Face-identification task (R^2 adjusted = 27.9%)				
Intercept	0.52928	0.09480	5.580	.000
Congruency × Alignment (d')	0.01139	0.01560	0.731	.470
Congruency × Alignment (RT)	0.00043	0.00010	3.030	.005
Misaligned faces (d')	0.05529	0.02210	2.500	.018
Misaligned faces (RT)	0.00014	0.00010	1.370	.179
Cambridge Face Memory Test (R^2 adjusted = 20.5%)				
Intercept	48.02460	17.80000	2.700	.011
Congruency × Alignment (d')	8.00504	2.92400	2.740	.010
Congruency × Alignment (RT)	0.04713	0.02640	1.790	.083
Misaligned faces (d')	5.81439	4.14700	1.400	.170
Misaligned faces (RT)	0.00397	0.01930	0.206	.838

Note: The only predictors that were significantly correlated were d' and reaction time (RT) for misaligned faces ($r = -.37$, $p = .02$).

Cottrell, 1999; Le Grand et al., 2006; Schwartz et al., 2002). If holistic processing is not predictive of face-identification performance (Konar et al., 2010), this undermines the motivation for this line of research and may require researchers to rethink how face recognition is studied. Fortunately, this is not necessary: Although there are some differences between the two face-processing tasks we used, individual differences in both face matching and face identification were related to holistic processing. However, the choice of holistic-processing measure seems to be critical: Our results were consistent with Konar et al.'s (2010) findings when the partial-design measure of holistic processing was used, but opposite conclusions were reached using a different measure in the same task.

Our confirmation of the relationship between holistic processing and face recognition corroborates a widely held assumption, but, counterintuitively, it is problematic in other ways. For example, we would expect that face-recognition abilities improve over the course of development, yet there are reports of young children exhibiting adultlike holistic processing (e.g., Macchi Cassia et al., 2009; Mondloch et al., 2007). Similarly, individuals with developmental prosopagnosia and patients with schizophrenia show deficits in face recognition. If holistic processing were critically related to face recognition, we would expect abnormal holistic processing for these groups. But in both cases, holistic processing is reported to be normal (Le Grand et al., 2006; Schwartz et al., 2002). However, these conclusions are based solely on research using the partial design, and they could reflect artifacts of important (and potentially informative) group differences in response biases. In another debate, holistic processing measured in the complete design is one of the hallmarks of face perception that can be acquired for nonface objects (Gauthier & Tarr, 2002; A.C.-N. Wong, Palmeri, & Gauthier, 2009), but partial-design studies have failed to replicate this result (Robbins & McKone, 2007).

Just as abandoning phrenology did not mean rejecting cortical specialization of functions, this is a case in which abandoning a flawed measure increases the construct validity of holistic processing. However, we cannot hope to make theoretical progress in our understanding of the mechanisms underlying face perception if we continue to use the partial design of the composite task. Holistic processing is a valuable construct that provides a link between experience, performance, and brain specialization. For instance, practice individuating objects produces increases in holistic processing that predict activity in the fusiform gyrus (A.C.-N. Wong, Palmeri, & Gauthier, 2009; A.C.-N. Wong, Palmeri, Rogers, Gore, & Gauthier, 2009). These previous studies offer experimental evidence for the causal influences of holistic processing on individuation ability that can only be inferred from the correlations obtained in the present study of face recognition.

Acknowledgments

We thank Tom Palmeri, Grit Herzmann, and Alan Wong for comments; Brad Duchaine for providing the Cambridge Face Memory Test; and Rankin McGugin and Sarah Muller for assistance with data collection and analyses.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by the James S. McDonnell Foundation, the National Science Foundation (Grant SBE-0542013), the Vanderbilt Vision Research Center (Grant P30-EY008126), and the National Eye Institute (Grant R01 EY013441-06A2).

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

References

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

Cheung, O.S., Richler, J.J., Palmeri, T.J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1327–1336.

Dailey, M.N., & Cottrell, G.W. (1999). Organization of face and object recognition in modular neural networks. *Neural Networks*, *12*, 1053–1074.

de Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. *Journal of Experimental Child Psychology*, *96*, 57–70.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic faces. *Neuropsychologica*, *44*, 576–585.

Farah, M.J., Wilson, K.D., Drain, M., & Tanaka, J.W. (1998). What is "special" about face perception? *Psychological Review*, *105*, 482–498.

Furl, N., Garrido, L., Dolan, R.J., Driver, J., & Duchaine, B. (2010). Fusiform gyrus selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*. Advance online publication. doi:10.1162/jocn.2010.21545

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

Gauthier, I., & Tarr, M.J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 431–466.

Goffaux, V. (2009). Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. *Vision Research*, *49*, 774–781.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"—Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1023–1039.

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

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

Le Grand, R., Cooper, P.A., Mondloch, C.J., Lewis, T.L., Sagiv, N., de Gelder, B., & Maurer, D. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition*, *61*, 139–158.

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

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

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

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

Mondloch, C.J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence for adult-like holistic processing. *Visual Cognition*, *15*, 564–577.

Richler, J.J., Bukach, C.M., & Gauthier, I. (2009). Context influences holistic processing of non-face objects in the composite task. *Attention, Perception, & Psychophysics*, *71*, 530–540.

Richler, J.J., Mack, M.L., Gauthier, I., & Palmeri, T.J. (2009). Holistic processing of faces happens at a glance. *Vision Research*, *49*, 2856–2861.

Richler, J.J., Mack, M.L., Palmeri, T.J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research*, *51*, 333–342.

Richler, J.J., Tanaka, J.W., Brown, D.D., & Gauthier, I. (2008). Why does selective attention to parts fail in face processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1356–1368.

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

Schwartz, B.L., Marvel, C.L., Drapalski, A., Rosse, R.B., & Deutsch, S.I. (2002). Configural processing in face recognition in schizophrenia. *Cognitive Neuropsychiatry*, *7*, 15–39.

Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schact, A., & Sommer, W. (in press). Individual differences in face cognition. *Journal of Personality and Social Psychology*.

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

Wong, A.C.-N., Palmeri, T.J., Rogers, B.P., Gore, J.C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PLoS ONE*, *4*(12), e8405. Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0008405

Wong, Y.K., & Gauthier, I. (2010). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience*, *22*, 695–713.

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

Supplemental Materials

Robustness of the Correlations. To verify the robustness of the significant correlations, we first searched for outliers using externally studentized residuals. At an alpha level of .05, we found only 1 outlier, and taking it out only increased the correlation. Second, we calculated Spearman rank order correlations, a measure known for its robustness and efficiency. All significant correlations remained significant (see Table 1).

Table 1.

	Pearson r	Without outliers (n=0,0,1)	95% confidence interval on r	Spearman rank order
CMFT with d'	r=.396, p=.014	same	.088 -> .635	r=.386,p=.02
CMFT with RT	r=.334, p=.04	same	.017 -> .590	r=.396,p=.01
Face ID with RT	r=.482, p=.002	r=.548,p=.0004	.274 -> .741	r=.461, p=.004

The partial design. In the partial design of the composite task (Figure 1), the relevant parts can be same or different, but the irrelevant parts are always different. Performance on these trials with aligned parts is compared to the same trials in a misaligned configuration. In the partial design, this other condition, sometimes misaligned and sometimes inverted faces, is necessary to provide an index of holistic processing which is performance in the aligned condition relative to a baseline.

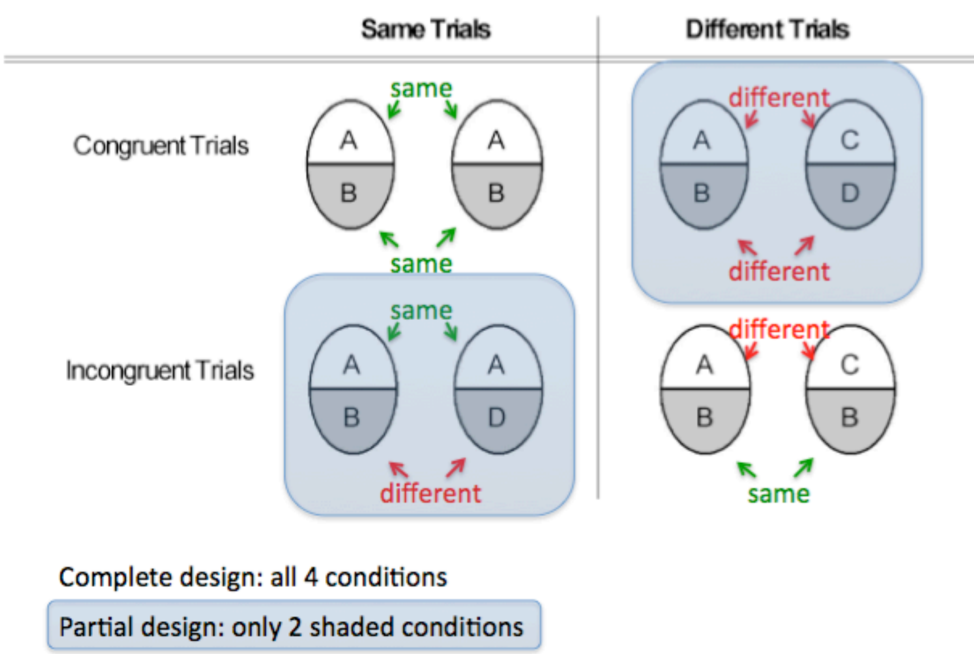

Figure 1. Conditions included in the complete and partial design of the composite task. Each pair of ovals represents two faces shown sequentially in a matching trial where participants are asked to match the relevant parts (in white) and ignore the irrelevant parts (in grey). Letters denote different identities.

It has been argued (correctly) that predictions can only be made about same trials (i.e., the different irrelevant parts should hurt performance on same trials if they cannot be ignored) and it is much harder to make a prediction about the different trials (because the B and D *different* bottoms could increase the difference that already exists for A and C tops, OR it could reduce it, depending on whether B is more or less similar to D than A is to C). For this reason, many authors analyze only accuracy for same trials in the partial design. Konar et al. (2010), used d' (using the hit rate on same trials, and false alarm rate on different trials). In the present work, we performed partial analyses using both hit rate and this d' index.

The complete design. The complete design includes the other half of the condition matrix, adding "same" and "different" trials with irrelevant parts that are the same. Note that regardless of condition composite faces are always made by combining top and bottom halves from different original faces. Accordingly, one can define "congruent" and "incongruent" trials, depending on whether the correct response corresponds to the response that would be given to the irrelevant part. In the complete design, predictions *can* be made regarding holistic processing for *both* same and different trials: if participants cannot selectively attend to the relevant part, they may be relatively impaired on incongruent relative to congruent trials. On same-incongruent trials, failure to selectively attend to the relevant parts could only make them more likely to respond different, and on different-incongruent trails, failure to selectively attend could only make them more likely to respond same. Misaligned (or inverted) trials are often included in the complete design to see if failures of selective attention are sensitive to configuration of the parts. Some work suggests that the most valid index of automatic holistic processing, sensitive to effects shown for faces and objects of expertise but not for objects in novices, is the congruency x alignment interaction (a congruency effect for aligned stimuli which is reduced for misaligned stimuli – see Richler et al., 2008b; Richler et al., 2009a; Richler et al., in press a). Therefore, the alignment effect in partial analyses within the complete design is compared to an alignment x congruency effect in complete analyses.

The response bias confound. The main problem raised with the partial design is that all same trials are incongruent and all different trials are congruent. Thus, the correct response is confounded with congruency. While congruency is not a factor of interest in the partial design, it is impossible to demonstrate that congruency does not influence response bias as it does in the complete design. Many studies (Cheung et al., 2008; Richler et al., 2008a; 2009b; in press a) show that participants are biased to say "different" for incongruent vs. to congruent trials. In the partial design, so it is impossible to separate the tendency to say "different" on same-incongruent trials from the tendency to say "different" on any incongruent trial regardless of the correct response.

In the complete design, where it is possible to measure response bias as a function of congruency, response bias often depends on alignment (Cheung et al., 2008; Richler et al., 2008a,b). This interaction can also be complicated either by stimulus manipulations (e.g., spatial frequency filtering; Cheung et al., 2008) or even simply by telling participants that there are more or less "same" trials (Richler et al., submitted). These response biases could be influenced by other task or group differences.

Critically, the problem is not solved simply by using signal detection measures. For instance, Konar et al. (2010) compute d' using accuracy on same-incongruent trials as hit rate and errors on different-congruent trials as false alarm rate. Even if biases related to alignment are taken into consideration by using this d', biases related to congruency that often differ between aligned and misaligned trials (Cheung et al., 2008; Richler et al., 2008a,b; in press a) remain confounded in these analyses (as demonstrated by the correlation between the partial measure alignment effect d' and the complete design measure of bias in the main paper).

Practically speaking, the advantage of the complete over the partial design does not end with the use of a d' measure that removes the congruency x alignment response bias, because the congruency x alignment effect in mean correct response times also provides a measure of holistic processing that correlates with face recognition, whereas the alignment effect in RTs does not (either in our study, or in the version by Konar et al., 2010). This is the case even though the partial alignment effect in RTs does not correlate with response bias in the complete design (r =-.112, n.s.) the way that the partial alignment effect in d' does. Response times can be subject to different criteria than accuracy and future work could improve the RT index of holistic processing using extensions of signal detection theory (e.g., Ratcliff, 1978; Balakrishnan et al., 2002).

The two designs in the literature. We argue that it is innapropriate to combine studies that use the partial and complete designs in any review or meta-analysis, because the response bias confounds are potentially serious enough that partial design results cannot be attributed solely or even mainly to holistic processing. To help clarify which composite studies used which design, Table 2 lists partial and complete design composite studies.

Table 2.

Partial Design Studies			**Complete Design Studies**		
Year	**Authors**	**Subject/Area**	**Year**	**Authors**	**Subject/Area**
1987	Young, Hellawell & Hay	inversion	1998	Farah, Wilson, Drain & Tanaka	faces vs. objects
1994	Hole	unfamiliar faces	2002	Gauthier & Tarr	expertise (objects)
	Carey & Diamond	developmental		Wenger & Ingvalson	decisional factors
1999	Hole, George & Dunsmore	contrast inversion	2003	Wenger & Ingvalson	decisional factors
2000	Calder, Young, Keane & Dean	facial expression		Gauthier, Curran, Curby & Collins	ERP; interference
2002	Boutet, Gentes-Hawn & Chaudhuri	attention	2005	Gauthier & Curby	interference
	Mondloch, Le Grand & Maurer	developmental	2006	Bukach, Bub, Gauthier & Tarr	prosopagnosia
	Schwartz, Marvel, Drapalski, Rosse & Deutch	Schizophrenia	2008	Cheung, Richler, Palmeri & Gauthier	spatial frequency
2003	Pellicano & Rhodes	developmental		Richler, Gauthier, Wenger & Palmeri	decisional factors
	Robbins & McKone	rotation		Richler, Tanaka, Brown & Gauthier	decisional factors/attention
	Teunisse & de Gelder	Autism	2009	Gauthier, Klaiman & Schultz	Autism
2004	Le Grand, Mondloch, Maurer & Brent	experience (faces)		Goffaux	spatial frequency
2005	Calder & Jansen	facial expression		Hsiao & Cottrell	expertise (chinese characters)
	Weston & Perfect	Global/local bias		Richler, Bukach & Gauthier	objects
2006	Khurana, Carter, Watanabe & Nijhawan	temporal integration		Richler, Mack, Palmeri & Gauthier	encoding time
	Singer & Sheinberg	temporal integration		Wong, Palmeri & Gauthier	expertise (objects)
	Michel, Caldara & Rossion	other-race effect	2010	Cheung & Gauthier	interference
	Michel, Rossion, Han, Chung & Caldara	other-race effect		Wong & Gauthier	expertise (objects)
	Schiltz & Rossion	fMRI		Todorov, Loehr & Oosterhof	social
	Goffaux & Rossion	spatial frequency		Bukach, Philips & Gauthier	expertise (objects)
	Le Grand, Cooper, Mondloch, Lewis, Sagiv, de Gelder & Maurer	developmental prosopagnosia		Richler, Mack, Palmeri & Gauthier	inversion effect
	Parr, Heintz & Akamagwuna	monkeys		Richler, Cheung & Gauthier	top-down influences
2007	Anaki, Boyd & Moscovitch	temporal integration		Richler, Cheung & Gauthier	individual differences
	De Heering, Houthuys & Rossion	developmental			
	Mondloch, Pathman, Maurer, Le Grand & Schonen	developmental			

	McKone, Brewer & MacPherson	other-race effect
	Michel, Corneille & Rossion	other-race effect
	Robbins & McKone	expertise (objects)
	Durand, Gallay, Seigneruic, Robichon & Baudouin	facial expression
2008	Abbas & Duchaine	attractiveness
	De Heering, Rossion, Turati & Simion	eye movements
	De Heering & Rossion	experience (faces)
	McKone	viewpoint
	Mondloch & Maurer	orientation
	Rossion & Boremanse	orientation
	Letourneau & Mitchell	ERP
	Nishimura, Rutherford & Maurer	Autism
	Hertzmann, Danthir, Schact, Sommer & Wilhelm	individual differences
2009	Taubert & Alais	
	Cassia, Picozzi, Kuefner, Bricolo & Turati	developmental
	Susilo, Crookes, McKone & Turner	experience (faces)
	Hugenberg & Corneille	other-race effect
	Michel, Corneille & Rossion	other-race effect
	Taubert	monkeys
	Jacques & Rossion	ERP
2010	Konar, Bennett & Sekuler	individual differences
	Zhao & Hayward	gender
	Ramon, Busigny & Rossion	prosopagnosia
	Mondloch, Elms, Maurer, Rhodes, Hayward, Tanaka & Zhou	other-race effect
	Zhu, Song, Hu, Li, Tian, Zhen, Dong & Kanwisher	individual differences
	Schiltz, Dricot, Goebel & Rossion	fMRI
	Kuefner, Jacques, Prieto & Rossion	ERP
	Kuefner, Cassia, Vescovo & PIcozzi	experience (faces)
	Wilhelm, Herzmann, Kunina, Danthiir, Schacht & Sommer	individual differences
	Taubert	monkeys

The configuration of the first face in each trial. Another notable difference between the composite design we ran here and the one in Konar et al. (2010) is that the first stimulus in each of our trials was always aligned (as in most studies using the complete design), whereas in Konar et al., the first and second stimuli were the same format (as in most studies using the partial design). We chose to use only aligned study stimuli because of empirical evidence that when the first stimulus is misaligned on some of the trials, a congruency effect can be contextually induced for non-face objects (Richler et al., 2009a). In contrast, when the first stimulus is always aligned, the congruency effect – and the congruency x alignment interaction – is not observed for objects in novices and is only obtained for faces or objects of expertise. This choice was therefore made to maximize the likelihood that we would tap into variance that is associated with face-selective holistic processing. Given evidence of contextual congruency effects for objects in novices and not for faces in experts, it is possible that misaligning faces at study would induce a contextual congruency effect especially in individuals with the poorest face recognition abilities, thereby further muddying the measurement of individual differences. We do not know of a systematic exploration of the importance of the format of the first face in the partial design. However, our partial analyses in a design where the first face is always misaligned produce no more or less evidence for a correlation with face recognition than the partial design where aligned-aligned trials were compared to misaligned-misaligned trials (Konar et al. 2010).

Partial design vs. partial analyses. We call the design where only the irrelevant part-different are included "partial design" – these conditions constitute half of the conditions included (and randomized) in the complete design (see Figure 1) and therefore it is possible to conduct "partial analyses" using only the relevant conditions. Because it is impossible to isolate such response biases using only the partial design trials, it is impossible to demonstrate that the response bias present in partial analyses in the complete design (using only shaded trials in Figure 1) are not present when these trials are collected in isolation. It is possible, although it cannot be tested, that using the partial design by itself does not lead to the same response biases as partial analyses of some of the trials in the complete design. However, this makes context yet another factor that may influence response bias in the partial analyses.

There might be a concern that the partial design trials cannot reflect the alignment effect to the same extent when they are acquired in the context of the other half of the trials in the complete design. The partial analyses measures in the main experiment do not yield a significant alignment effect at the group level (and the alignment effect is highly variable, as it is in Konar et al., 2010). It is possible that when misaligned trials use a misaligned face at study (which can inflate the congruency effect as shown in Richler et al., 2009a) this also affects bias in the partial design and impacts the alignment effect. Because such conjectures cannot be tested in the partial design, we provide here some empirical evidence that the partial measures obtained within the complete design can in principle capture the same alignment effect seen in the partial design, at least when misaligned trials use misaligned faces at study. The fact that we do not observe a mean alignment effect here is likely due to the poor reliability of this measure rather than the specific design we used.

To make this point, we compare data from a prior study where we used similar methods (complete design of the composite task, comparing complete and partial analyses; Cheung et al., 2008), and unpublished comparison data that had been collected to address this concern, from another group of participants who performed the same partial design trials on their own.

Methods

The details of the methods are provided in Cheung et al., (2008). The experiment used images of 20 faces that were either low-pass filtered (LSF), high-pass filtered (HSF) or not filtered (full-spectrum, FS). Top and bottom halves of these images were randomly paired from different individuals of the same sex. A study composite face was shown for 600 ms, followed by a 300-ms blank, followed by a test composite face for 1 s or until a response was made (whichever came first).

The only difference between the two groups was that the 240 partial design trials (80 trials for each of the spatial frequency conditions, including 40 same and 40 different trials) were either run randomized together with the other half of the trials (Complete context – data reported in Cheung et al., 2008, n= 21) or by themselves (Partial only context– unpublished data, n=24).

Results

A 2 (context: complete context vs. partial context) × 3 (spatial frequency: LSF vs. HSF vs. FS) × 2 (alignment: aligned vs. misaligned) ANOVA was conducted on both accuracy and RT for "same" trials in the partial design. There were no significant main effects of Context (accuracy: $F(1,43)=.06$, $p>.806$; RT: $F(1,43)=.08$, $p>.778$). There were also no interactions with Context and any other factors in accuracy (F's< .64, p's>.53) or RT (F's<.58, p's>.45). Therefore, the results from the partial design were unaffected by whether or not these trials were presented on their own or in the context of the complete design (see Figure 2).

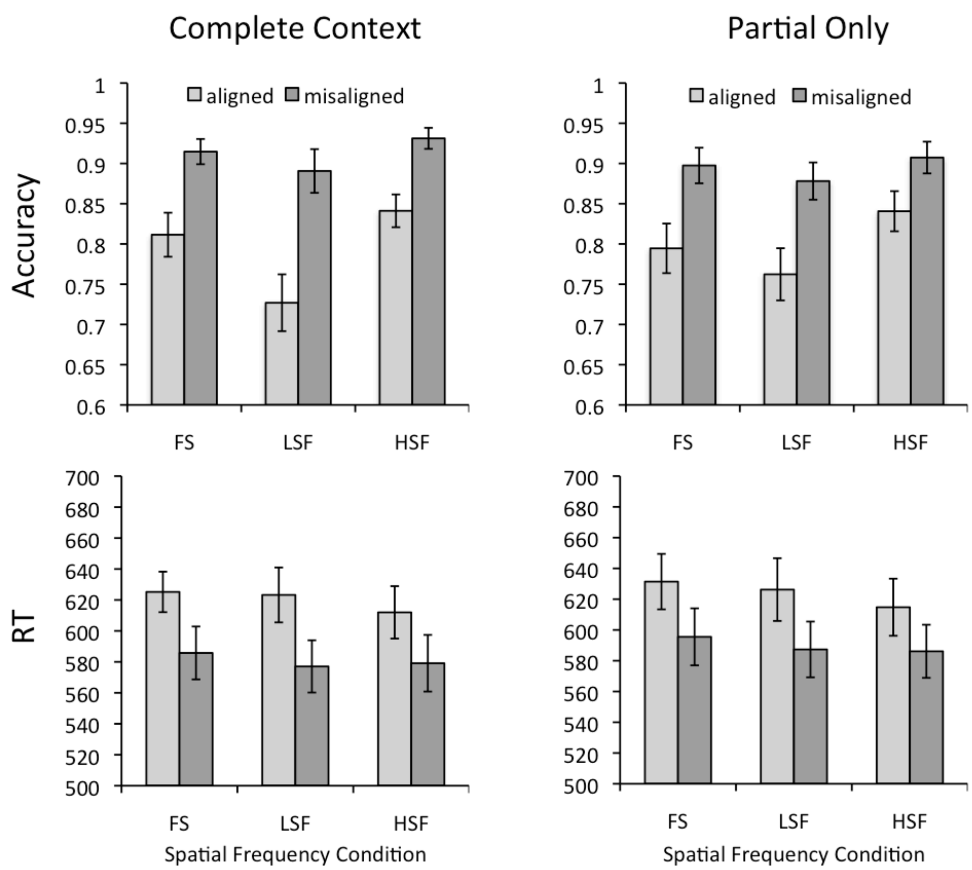

Figure 2. Accuracy (top row) and RT (bottom row) for "same" trials in the partial design as a function of spatial frequency content when the partial design trials are presented in the context of the complete design (left) or by themselves (right). Error bars show standard error of the mean.

References

Balakrishnan, J.D., MacDonald, J.A., Busemeyer, J.R. & Lin, Anli. (2002). Dynamic Signal Detection Theory: The next logical step in the evolution of signal detection analysis. Technical Report #248, Indiana University Cognitive Science Program.

Cheung, O.S., Richler, J.J., Palmeri, T.J. & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception & Performance, 34,* 1327-1336.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59-108.

Richler, J.J., Bukach, C.M., & Gauthier, I. (2009a). Context influences holistic processing of non-face objects in the composite task. *Attention, Perception & Psychophysics, 71,* 530-540

Richler, J.J., Cheung, O.S., & Gauthier, I. (submitted). Beliefs alter holistic face processing…if response bias is not taken into account.

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

Richler, J.J., Mack, M.L., Gauthier, I., & Palmeri, T.J. (2009b) Holistic processing of faces happens at a glance. *Vision Research, 49,* 2856-2861.

Richler, J.J., Mack, M.L., Palmeri, T.J., & Gauthier, I. (in press a). Inverted faces are (eventually) processed holistically. *Vision Research.*

Richler, J.J., Tanaka, J.W., Brown, D.D. & Gauthier, I. (2008b). Why does selective attention to parts fail in face processing? *Journal of Experimental Psychology: Learning, Memory and Cognition, 34,* 1356-1368.

Richler, J.J., Wong, Y.K., & Gauthier, I. (in press b). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science.*