Holistic face perception is modulated by experience-dependent perceptual grouping

Kim M. Curby^{1,2} · Robert J. Entenman³ · Justin T. Fleming⁴

Published online: 30 March 2016
© The Psychonomic Society, Inc. 2016

Abstract What role do general-purpose, experience-sensitive perceptual mechanisms play in producing characteristic features of face perception? We previously demonstrated that different-colored, misaligned framing backgrounds, designed to disrupt perceptual grouping of face parts appearing upon them, disrupt holistic face perception. In the current experiments, a similar part-judgment task with composite faces was performed: face parts appeared in either misaligned, different-colored rectangles or aligned, same-colored rectangles. To investigate whether experience can shape impacts of perceptual grouping on holistic face perception, a pre-task fostered the perception of either (a) the misaligned, differently colored rectangle frames as parts of a single, multicolored polygon or (b) the aligned, same-colored rectangle frames as a single square shape. Faces appearing in the misaligned, differently colored rectangles were processed more holistically by those in the polygon-, compared with the square-, pre-task group. Holistic effects for faces appearing in aligned, same-colored rectangles

showed the opposite pattern. Experiment 2, which included a pre-task condition fostering the perception of the aligned, same-colored frames as pairs of independent rectangles, provided converging evidence that experience can modulate impacts of perceptual grouping on holistic face perception. These results are surprising given the proposed impenetrability of holistic face perception and provide insights into the elusive mechanisms underlying holistic perception.

Keywords Face perception · Object-based attention · Object recognition

Electronic supplementary material The online version of this article (doi:10.3758/s13414-016-1077-8) contains supplementary material, which is available to authorized users.

✉ Kim M. Curby
kim.curby@mq.edu.au

¹ Department of Psychology, Macquarie University,
NSW 2109 Sydney, Australia

² Center for Elite Performance, Expertise, & Training, Macquarie University, Sydney, Australia

³ Department of Psychology, Temple University, Philadelphia, PA, USA

⁴ Division of Medical Sciences, Harvard University, Cambridge, MA, USA

To those who study face perception, an intriguing question over the past several decades has been: how and why are faces perceived as being so much more than the sum of their parts? While there has been much progress in understanding the intricacies of behavioral markers that distinguish face from non-face object perception (Calder, Rhodes, & Johnson, 2011; Richler & Gauthier, 2014), a full mechanistic understanding of face perception is still elusive. Perceptual tasks comparing face and non-face object perception often reveal performance differences, and such differences often are interpreted as evidence that face perception involves the recruitment of *qualitatively* distinct mechanisms. However, it is possible that more general-purpose perceptual mechanisms, shaped by experience, contribute to the characteristic features of face perception. We investigated the possibility that holistic face perception effects are supported by perceptual grouping mechanisms that more generally underlie the perception of objecthood.

Facial features appear to be obligatorily processed in the context of their relations with other facial features, yielding impressive sensitivity to subtle changes in the configuration of these features (Bruce, Doyle, Dench, & Burton, 1991; Haig,

1984) but also rendering it difficult to selectively focus on some facial features while ignoring others (Gauthier, Curran, Curby, & Collins, 2003; Richler, Tanaka, Brown, & Gauthier, 2008; Young, Hellawell, & Hay, 1987). This has been referred to as “holistic” processing and contrasts with the more “feature-based” processing typical of non-face object processing. Numerous studies have highlighted the effects of experience on the development of face- and non-face-related holistic perceptual styles. For example, some studies have reported evidence that other-race faces, for which observers typically have less experience with, are processed less holistically than are same-race faces (Michel, Caldara, & Rossion, 2006; Michel, Rossion, Han, Chung, & Caldara, 2006; Tanaka, Kiefer, & Bukach, 2004), and individual differences in exposure to other-race faces were found to predict the degree to which they are processed holistically (Bukach, Cottle, Ubiwa, & Miller, 2012; Tanaka et al., 2004 but see Zhao, Hayward, & Bülthoff, 2014 for evidence suggesting a more complex relationship between other-race effects and experience). Furthermore, markers of holistic perception have been found among experts with non-face categories, such as cars, fingerprints, and chess configurations (Boggan, Bartlett, & Krawczyk, 2012; Busey & Vanderkolk, 2005; Curby, Glazek, & Gauthier, 2009; Gauthier et al., 2003). One intriguing possibility is that holistic aspects of face perception may, in part, reflect experience-based strengthening of perceptual grouping of features within faces, rendering the features of faces into more coherent wholes relative to non-face objects.

Strengthening the perceptual grouping of facial features would serve to fortify the *objecthood* of the face. *Objecthood* refers to the perception of a shared structure and/or identity across visual features or parts (Palmer, 1977) and is influenced both by image-based cues within a stimulus (Palmer, 1977) and observers’ experience with the object (Vecera & Farah, 1997). The perception and strength of the objecthood of a given stimulus can dramatically impact its perception, leading to many perceptual benefits.

A prominent example of such perceptual benefits comes from the literature on *object-based attention*, which refers to instances when objects become the unit of selection (in contrast to when a spatial location or stimulus feature, e.g., the color red, is the basis of selection). In such cases, when attention is directed to part of an object, other parts of that object enjoy a processing advantage (Egley, Driver, & Rafal, 1994). Furthermore, the allocation of attention across features is more efficient when the features are perceived as being part of the same object.

Indexes of object-based attention are strongly influenced by manipulations of perceptual grouping cues designed to impact the perception of objecthood (Matsukura & Vecera, 2006), as well as observers’ previous experience with the stimulus (Vecera & Farah, 1997; Zemel, Mozer, Behrmann, & Bavelier, 2002). For example, components of upright letters

appear to be more strongly grouped than those from inverted letters, or non-familiar shapes (Kimchi & Hadad, 2002; Vecera & Farah, 1997; Vickery & Jiang, 2009). In addition, participants are faster to make a judgment about two features on separate fragments behind an occluding shape when their experience is consistent with these two features belonging to the same object versus different objects (Zemel et al., 2002). Thus, strengthening a stimulus’ perceived objecthood strengthens the degree to which it is attended as a singular perceptual unit. The full potential and consequences of experience-based changes to mechanisms underlying the perception of objecthood, such as in the case of faces, is relatively unknown.

Object-based attention effects can also be a double-edged sword; facilitated perception for features within an attended object occurs even when such features are task-irrelevant and, as a result, object-based attention can produce performance *costs* when such features contain misleading information. For example, flanker interference is greater when non-targets flanking a target are perceived to be on the same object (Kramer & Jacobson, 1991). The automatic spreading of attention to (Hollingworth, Maxcey-Richard, & Vecera, 2012; Richard, Lee, & Vecera, 2008), or attentional prioritization of (Shomstein & Yantis, 2002), disparate locations within a cued object can effectively produce a failure of selective attention.

Intriguingly, some markers of holistic face perception have been characterized as failures of selective attention in terms of selectively attending to either specific types of information (configural or componential; Amishav & Kimchi, 2010) or spatial regions (top or bottom; Richler et al., 2008). For example, in the composite face task, where observers make judgments about only the top or bottom halves of faces, this failure takes the form of an inability to selectively attend to the task-relevant face region. As a result, observers experience interference from the other, task-irrelevant halves. Holistic perception is indexed as the degree to which manipulations of the task-irrelevant part impact judgments involving only the task-relevant part. Towards this end, holistic processing can be indexed via the *composite effect*: the cost in accuracy and/or response time to identifying part of a face (e.g., the top half) when it is *aligned* with the bottom half of a different face, relative to when these parts are *misaligned* (which disrupts holistic perception). Part-matching judgments about the tops or bottoms of composite faces also are impacted by the compatibility of the relationship between the task-irrelevant face parts. Specifically, judgments about the task-relevant (cued) parts are impacted by whether the relationship between the task-irrelevant parts is the same (congruent) or different (incongruent) to that between the task-relevant parts. Face part judgment performance is poorer when the relationship between the task-irrelevant parts is incongruent, compared with congruent, with that between the task-relevant parts. This effect forms the basis of another index of holistic processing

known as the *congruency effect* (Curby, Johnson, & Tyson, 2012; Gauthier et al., 2003; Richler et al., 2008). The modulation of the congruency effect by the alignment of the composite face parts is suggested to be a key feature of holistic processing (Richler, Bukach, & Gauthier, 2009). In sum, holistic perception of faces can be characterized as the processing of a face as a singular perceptual unit and difficulty limiting attention to a specific subregion or part of the face.

Although the disruptive effect of misalignment on holistic processing—whereby it enables participants to more effectively selectively attend to only part of the face—is important for isolating holistic perception, it is unclear how or why misalignment wields its impact. Of note is that misalignment disrupts the Gestalt grouping cue of good continuation of the external and internal face contours. Thus, one possibility is that misalignment disrupts holistic face perception via its impact on weakening the perceptual grouping of the face parts, and thus the shared identity or objecthood of the face. This would result in weakened object-based attention and thus would better allow observers to constrain their attention to only the task relevant part. Thus, like in the case of documented object-based attention effects, when the face parts are aligned, attention may obligatorily spread throughout the face, resulting in a task-irrelevant face region impairing judgments about other regions of the face.

Recent evidence supports the potential of basic perceptual mechanisms to contribute to characteristic holistic perception effects observed for faces, specifically the disruption to holistic perception with misalignment. We recently reported that holistic perception, indexed using the composite face paradigm, was disrupted by contextual cues that discouraged the perceptual grouping of face parts (Curby, Goldstein, & Blacker, 2013). In that study the top and bottom halves of composite faces appeared in either 1) aligned, same-colored frames, which presumably encouraged the perceptual grouping of the two parts or 2) misaligned, differently-colored rectangular frames, which would have discouraged the grouping of the parts into a single perceptual unit. In both cases, the top and bottom halves of the face images were always aligned. Part judgments for faces appearing in frames that discouraged perceptual grouping showed less interference from the task irrelevant face part in the form of a reduced congruency effect compared with faces appearing in frames encouraging the perceptual grouping of the two face parts. The ability of such perceptual grouping cues to impact holistic face perception even when the face parts themselves are aligned is particularly noteworthy. Critically, this effect was face- and orientation-specific, with the perception of cars (amongst car novices) or inverted faces failing to show the same vulnerability to the disruption of perceptual grouping mechanisms.

We suggest that holistic face perception, as indexed via the composite face effect, may stem in part from the allocation of attention to the face as a unit (i.e., object-based attention), to

both the task-relevant and -irrelevant face parts (Curby et al., 2013). Physically misaligning the face parts, as in the classic composite face effect, disrupts the cohesiveness of the face as a unit of selection, thereby allowing attention to more effectively target the task-relevant part. Consistent with this possibility, other manipulations known to disrupt objecthood, and thus object-based attention, such as spatially separating different parts of an object or presenting them on different depth planes, also have been shown to disrupt holistic face perception (Taubert & Alais, 2009).

Experiment 1

If holistic face perception is supported by experience-based strengthening of the “objecthood” or cohesion of the face as a perceptual unit, then the disruptive effect of perceptual cues discouraging the grouping of face parts on holistic perception of the faces (previously reported in Curby et al., 2013) also should be shapeable by an observer’s prior experience. In Experiment 1, we tested this prediction by having participants perform the same part-judgment task with composite faces used previously (Curby et al., 2013) but after one of two pre-tasks. One pre-task encouraged observers to perceive the misaligned, differently colored rectangular shapes (which the face composites subsequently appeared in) as a single, irregular polygon unit. The other, control pre-task, reinforced the perception of the aligned, uniformly-colored rectangles as a single square-like shape, consistent with the low-level features driving their perceptual grouping. If experience-dependent object-based attention influences holistic perception of faces, when the face parts appear in polygon frames, observers who had performed the pre-task encouraging grouping of the rectangles into polygons should show a greater congruency effect than those who had performed the control pre-task. In contrast, when faces appear in the square framing shape, those observers who had performed the control pre-task encouraging grouping of the rectangles into a single square should show a similar or greater congruency effect than those who had performed the pre-task with polygons. (A greater congruency effect among the control pre-task group in this condition would suggest that the pre-task with squares boosted perceptual grouping beyond that supported by perceptual grouping features of the aligned, uniformly-colored rectangle pairs, i.e., the Gestalt cues of good continuation of contours and [color] similarity).

Method

Participants Fifty-nine undergraduate psychology students (9 males; mean age = 20.7 years, standard deviation [SD] = 6.1 years) from Macquarie University participated

for course credit.¹ All participants reported having normal or corrected-to-normal vision. Thirty were randomly assigned to the polygon training condition (4 males; mean age = 20.5 years, SD = 5.9 years) and 29 to the (control) square training condition (5 males; mean age = 20.8 years, SD = 6.4 years). Data from five participants in the polygon training group and seven in the square training group were not included in the final sample due to poor performance (see the [Results](#) section for details), leaving a final sample of 47 participants (25 in the polygon training group, 22 in the square training group).

Stimuli The stimuli used in the training phase consisted of a series of blue and red square-like shapes in three sizes ($2.0^\circ \times 3.34^\circ$, $3.15^\circ \times 3.25^\circ$, $4.0^\circ \times 3.25^\circ$) and a series of irregular polygons also in three sizes ($2.72^\circ \times 3.34^\circ$, $3.87^\circ \times 3.25^\circ$, $4.72^\circ \times 3.25^\circ$). The polygons were made from two conjoined, identically size rectangles, which were different in color (one red, one blue) and stacked vertically in such a way that the top was misaligned to the right (by 0.72°). The squares shapes were also made from stacking the same rectangular fragments, but they were identical in color and were aligned (see Fig. 1 for sample stimuli). The stimuli in the test phase consisted of twenty grayscale front-view images of male faces wearing neutral expressions (Meissner, Brigham, & Butz, 2005), cropped in an oval to remove the hair and other external features. These faces ($4.9^\circ \times 6.5^\circ$) were then divided into two equal halves along a horizontal plane aligned with the bridge of the nose. These face part stimuli were placed on red or blue rectangles ($7.5^\circ \times 4.0^\circ$). Three versions of these stimuli were created: one with the face part centered on the colored rectangle, and two in which the rectangle (but not the face part) was shifted to the left (top part) or right (bottom part) by approximately 1° . When the top and bottom parts are combined this produced a total frame (rectangle) misalignment for the composite faces of 2° (see Fig. 2 for sample stimuli). A black line of 0.2° thickness appeared over the seam between the two face parts (and the rectangles in which they appeared).

Design and Procedure Participants placed their chin on a chinrest that was 60 cm from a monitor (23.5" viewing area) aligned with the central fixation cross. The experimental task was built and run using E-Prime (Psychology Software Tools, Pittsburgh, PA).

Training Pre-Task During the training phase, participants performed a visual search task. Each trial started with a central fixation cross (50 ms), followed by a small, medium, or large white target shape (2500 ms). The search display was then presented and remained for 4000 ms or until the participant responded.

There were two training conditions: (A) Participants in the single (square) shape training group searched for a small, medium, or large, red or blue square that matched the size of the target shape, among an array of 12 such square shapes varying in size (Fig. 1a). (b) Participants in the compound (polygon) shape training group searched for a dual colored conjoined rectangle pair (forming a polygon), matching the size of the target shape, from among an array of 12 such polygons varying in size (Fig. 1b).

Each stimulus in the search array was presented in a cell in an invisible 4×4 grid. The exact location of the shape within the cell was randomly jittered by approximately 0.15° left, right, up, or down across trials. Participants' indicated via a key press whether a shape identical in size to the white target shape was or was not present in the search display. Participants were required to perform a minimum of 12 practice trials and could start the experiment after achieving $>75\%$ accuracy across the last 12 trials. Participants then performed six blocks of 20 such trials for a total of 120 trials, with a rest offered in-between each block. The target shape was present in the search display on 50 % of trials. During a 1000-ms intertrial interval, participants were given feedback if they responded incorrectly.

Test Phase During the test phase participants performed a modified version of the composite face task similar to that described previously (Curby et al., 2013). Specifically, face composites were created online for each trial by combining the top of one face with the bottom of another face. The top and bottom parts were presented in rectangular frames that were either 1) aligned and the same color (both red or both blue), or 2) misaligned (by 2°) and of different colors (one red, one blue). Crucially, in all cases the two face parts were aligned to form an intact face, with the alignment manipulation being limited to the framing rectangles (Fig. 2). Each trial started with a central fixation cross (1000 ms), followed by a composite face comprising of the top and bottom parts of different faces (1000 ms). This was masked by a 1000-ms textured pattern that also contained either a bracket around the top or bottom part indicating which part was task-relevant. Following the mask, a second face and the bracket cue were briefly presented (200 ms), followed again by the textured pattern and bracket cue (for 2400 ms or until response). Participants indicated via a key press whether the cued half (top or bottom) of the second face was the same as or different from the first face. As in the visual search task, feedback was provided during a 1000 ms intertrial interval. Participants

¹ Despite the order of presentation in this report, Experiment 2 was conducted before Experiment 1. Thus, the effect size of the group $\times$ congruency $\times$ alignment interaction in Experiment 2 provided the effect size for a power analysis to determine the required sample size to achieve 95 % power in Experiment 1. This analysis indicated that 48 participants would be required. However, assuming data loss would be similar to that in Experiment 2 (10 participants), data from 59 participants was collected to ensure sufficient power.

Fig. 1 Trial structure used for the shape visual search pre-task, with examples from the (a) square (aligned rectangle pairs) and (b) polygon (misaligned rectangle pairs) search conditions. The target shape was

always presented as a silhouette and participants were instructed to search for an item in the array that had the same global shape regardless of the colors of the internal parts

completed four blocks of 32 trials containing faces in either aligned or misaligned rectangle frames, with the order of blocks counterbalanced across participants. Within each block, for half of the trials, the same/different relationship between the task irrelevant (noncued) face parts in the two faces was *congruent* with the relationship between the task-relevant (cued) parts. In other words, if the task-relevant parts differed between the two faces, thus rendering the correct response for the trial “different,” the task-irrelevant face parts also differed from each other in congruent trials. In the other half of the trials within a block, the same/different relationship between the task irrelevant (noncued) face parts in the two faces was *incongruent* with the relationship between the task-relevant (cued) parts. For example, if the task-relevant parts differed between the two faces, thus again rendering

the correct response for the trial “different,” the task-irrelevant face parts were the same as each other. Participants completed a total of 128 face trials, consisting of 64 congruent and 64 incongruent trials.

Results

Five participants from the polygon visual search training group and seven from the square training group were excluded due to poor performance on the training task (mean accuracy < 65 %; 2 participants) or test task (mean d' < 1; 9 participants or negative d' in at least one condition; 1 participant). Visual search training task trials with response times greater than 3000 ms were excluded from analyses.

Fig. 2 Trial structure used for the modified composite face task. The face parts either appeared in different colored, misaligned rectangles (top) or same colored, either both red or both blue, aligned rectangles (bottom). A

dashed bracket served as the cue in each trial to indicate which face part (top or bottom) the participant should make a same/different judgment on

Visual Search (Training Task) A one-way ANOVA comparing accuracy on the visual search task containing squares with that containing polygons revealed a main effect of task, $F(1, 45) = 10.29$, $p = 0.003$, with higher accuracy in the square condition ($M = 90.6\%$, $SD = 6.6\%$) than the polygon condition ($M = 82.3\%$, $SD = 10.3\%$). An additional 2 (stimuli: square vs. polygon) $\times 2$ (target presence: present vs. absent) ANOVA on the response time (RT) data revealed a main effect of target presence, with the classic pattern of faster response times for target present trials than target absent trials, $F(1, 45) = 144.35$, $p \leq 0.0001$, $\eta_p^2 = 0.76$. There also was a main effect of stimuli, $F(1, 45) = 0.13$, $p = 0.716$, $\eta_p^2 < 0.01$. However, this was qualified by an interaction between stimuli and target presence, $F(1, 45) = 9.04$, $p = 0.004$, $\eta_p^2 = 0.17$. Scheffé post-hoc comparisons revealed that while visual search RTs for target absent square ($M = 1593$ ms) and polygon ($M = 1541$ ms) trials were similar, $p < 0.161$, RTs for target present trials differed for square ($M = 1202$ ms) and polygon ($M = 1306$ ms) arrays ($p = 0.007$). Thus, when present, participants were slightly faster to locate the square relative to the polygon target. However, when targets were not present, participants spent a similar amount of time searching the square and polygon displays.

Composite Face Task (Test Phase)

Sensitivity (d') Data were submitted to a mixed between-subjects repeated measures 2 (training: square vs. polygon) $\times 2$ (frame: aligned (square) vs. misaligned (polygon)) $\times 2$ (congruency: congruent vs. incongruent) ANOVA, revealing a significant main effect of congruency, $F(1, 45) = 180.68$, $p \leq 0.0001$, $\eta_p^2 = 0.80$, but not of training, $F(1, 45) = 0.25$, $p = 0.620$, $\eta_p^2 < 0.01$, or frame, $F(1, 45) = 0.17$, $p = 0.896$, $\eta_p^2 < 0.01$. There was a trend for an interaction between congruency and frame, $F(1, 45) = 3.18$, $p = 0.081$, $\eta_p^2 = 0.07$, with the expected larger effect of congruency for faces in aligned than misaligned frames. There were no significant interactions between any of the variables (all $p > 0.17$; see Table 1 of Supplementary Material section).

Response Time A mixed between-subjects repeated measures 2 (training: rectangle vs. polygon) $\times 2$ (frame: aligned (square) vs. misaligned (polygon)) $\times 2$ (congruency: congruent vs. incongruent) ANOVA was conducted on the response time data. There was a significant main effect of congruency, $F(1, 45) = 33.61$, $p \leq 0.0001$, $\eta_p^2 = 0.43$, and of frame, $F(1, 45) = 4.12$, $p = 0.048$, $\eta_p^2 = 0.08$, but not of training, $F(1, 45) = 0.05$, $p = 0.830$, $\eta_p^2 < 0.01$. As in the sensitivity (d') analysis, there was a trend for an interaction between congruency and frame, $F(1, 45) = 3.17$, $p = 0.082$, $\eta_p^2 = 0.07$, with the expected larger effect of congruency for faces in aligned than misaligned frames (Fig. 3; $p > 0.52$ for all other two-way interactions). However, there was a significant three-way interaction

Fig. 3 Mean response time (RT) for the congruent (diamonds) and incongruent (squares) conditions, and the resulting index of holistic perception (congruency effect, filled bars, reflecting the difference between these conditions) for faces presented in aligned, same colored rectangles (left) or misaligned, different colored rectangles (right). Holistic perception of faces appearing in the misaligned, dual colored frames was greater among participants who had performed the pre-task encouraging perception of this frame as a single polygon unit compared with those who had performed the pre-task with squares. In contrast, holistic perception of faces appearing in the aligned, same colored frames was greater among participants who had performed the control pre-task with squares, compared to those who had performed the pre-task with polygons. Error bars represent standard error values

between congruency, frame, and training, $F(1, 45) = 7.22$, $p = 0.010$, $\eta_p^2 = 0.14$.

To investigate the basis of this three-way interaction, two additional two-way ANOVAs were conducted. Specifically, separate 2 (training: grouped vs. ungrouped) $\times 2$ (congruency: congruent vs. incongruent) ANOVAs were conducted on the RT data from the trials with the misaligned (polygon) and aligned frames. These analyses allowed us to specifically compare the impact of the different pre-task training conditions on the processing of the same visual stimulus. For the analysis of the data from the misaligned frame trials, there was a significant main effect of congruency, $F(1, 45) = 11.92$, $p = 0.001$, $\eta_p^2 = 0.21$, but no main effect of training, $F(1, 45) = 0.13$, $p = 0.723$, $\eta_p^2 < 0.01$. However, there was a significant two-way interaction between training and congruency, $F(1, 45) = 5.35$, $p = 0.025$, $\eta_p^2 = 0.11$, with participants who were trained to view the misaligned frames as a single polygon shape, showing a greater congruency effect (i.e., impact of the task-irrelevant part) than those in the control, square training, condition (Fig. 3). The 2 (training: grouped vs. ungrouped) $\times 2$ (congruency: congruent vs. incongruent) ANOVA conducted on the RT data from the trials with aligned (square) frames revealed a significant main effect of congruency, $F(1, 45) = 22.27$, $p \leq 0.0001$, $\eta_p^2 = 0.33$, but no main effect of training, $F(1, 45) = 0.01$, $p = 0.933$, $\eta_p^2 = 0.01$, and no interaction between training and congruency, $F(1, 45) = 2.58$, $p = 0.115$, $\eta_p^2 = 0.05$ (Fig. 3).

Discussion

Consistent with our previous work, holistic perception of faces was impacted by perceptual grouping cues that either encouraged or discouraged the grouping of face parts (Curby et al., 2013). Strikingly, we demonstrated that the previously documented disruption to the grouping of face parts appearing in misaligned, differently colored frames were attenuated in observers who had performed a pre-task that encouraged the perception of these misaligned, dual colored frames as a single polygon shape, relative to those who had performed the control (square) pre-task. Specifically, observers who performed a pre-task searching through arrays of polygons, formed by conjoined, misaligned and differently colored rectangle pairs, processed faces appearing in these shapes more holistically than those who had performed a similar pre-task with square shapes. These findings suggest that experience and learning modulate the impact of grouping cues on holistic face perception.

The above patterns of results were found in the response time but not the sensitivity (d') data. The sensitivity data showed a robust effect of congruency, but unlike the RT data, it was not impacted by pre-task training, nor the format of the framing rectangles. Thus, importantly, there was no evidence of a speed accuracy trade-off driving these findings.

Experiment 2

The results of Experiment 1 support the suggestion that experience-based grouping cues can impact the degree to which faces are processed holistically. As an additional test of this hypothesis, we also conducted an experiment in which a pre-task with the rectangle fragments themselves, instead of the squares made by conjoining them, were used, thereby encouraging observers who performed this pre-task to view the shapes framing the face parts as two independent rectangles, rather than a single square (aligned frame condition) or polygon (misaligned frame condition) shape. We included the pre-task with polygon shapes again for comparison. This also provided an opportunity to document the replicability of the facilitation of holistic processing of faces on misaligned, dual-colored rectangle fragments after observers were encouraged to perceive these elements as a singular polygon shape, as revealed in Experiment 1. As in Experiment 1, we predicted that observers who performed the polygon pre-task will show greater holistic processing of the faces appearing in the misaligned frames, but this time compared to observers who performed the pre-task with rectangles. In addition, we predicted that both pre-task groups will show similar levels of holistic processing for faces appearing in the aligned rectangle frames, as neither pre-task is expected to preferentially foster a single object-representation of the square frame. In contrast,

both pre-tasks may foster the perception of the aligned frames as separate elements (either as fragments from the pre-task polygons or as pairs of the pre-task rectangles), in conflict with the stimulus-based cues encouraging their grouping (i.e., their similar color and good continuation of the external contour).

Methods

Participants Seventy-six undergraduate psychology students (23 males; mean age = 20.3 years, SD = 2.4 years) from Temple University participated for course credit or monetary payment. All participants reported having normal or corrected to normal vision. Thirty-eight were randomly assigned to a polygon training condition (11 males; mean age = 20.4 years, SD = 2.8 years) and 38 to a rectangle training condition (12 males; mean age = 20.3 years, SD = 1.8 years). Adopting the same criteria as in Experiment 1, data from six participants from the polygon training group and four from the rectangle training group were not included in the final sample due to poor performance (see the Results section for details), leaving a final sample of 66 participants (32 in the polygon training group, 34 in the rectangle training group).

Stimuli In the training task, the stimuli were the same as those used in Experiment 1, except that the square shapes were replaced with the red and blue rectangle fragments that formed the polygon and square shapes. In the test phase, the stimuli (faces and their framing shapes) were the same as in the test phase in Experiment 1.

Design and Procedure Participants were seated approximately 60 cm from a monitor (21.5" viewing area) and again were required to rest their chin on a chinrest aligned with the central fixation cross.² Participants performed the same training and test phase as in Experiment 1, except, depending on what pre-task group they were randomly assigned to, participants either searched through arrays of the irregular polygons or just the red and blue rectangle shapes (the fragments used to form the polygon shapes). Rectangle arrays contained 24 items, whereas polygon arrays contained 12 items (as in Experiment 1), with each item presented within a randomly selected cell in an invisible 5×5 or 4×4 grid, respectively. Different set sizes in the two training conditions were selected to better match the number of "elements" (i.e., colored rectangles; each polygon contained two rectangle components) in each display.

² The use of a slightly smaller monitor, set at the same resolution, resulted in the stimuli appearing ~ 10 % smaller in Experiment 2 than in Experiment 1.

Results

Adopting the same criteria as in Experiment 1 resulted in the exclusion of six participants from the polygon visual search training group and four from the rectangle training group due to poor performance on the training task (mean accuracy < 65 %; 4 participants) or test task (mean d' < 1; 5 participants or negative d' in at least one condition; 1 participant). Response times greater than 3000 ms were excluded from analyses of the visual search training task.

Visual Search (Training Task) A one-way ANOVA comparing accuracy on the visual search task with rectangles ($M = 84.5\%$, $SD = 6.3\%$) with that with polygons ($M = 83.9\%$, $SD = 8.0\%$) revealed no significant difference, $F(1, 64) = 0.13$, $p = 0.72$, $\eta^2_p < 0.01$. An additional 2 (stimuli: rectangle vs. polygon) $\times 2$ (target presence: present vs. absent) ANOVA on the RT data revealed a main effect of target presence, with the classic pattern of faster response times for target present trials than target absent. In addition, there was a main effect of stimuli, $F(1, 64) = 6.37$, $p = 0.01$, $\eta^2_p = 0.09$, and an interaction between stimuli and target presence, $F(1, 64) = 21.11$, $p < 0.001$, $\eta^2_p = 0.25$. Scheffé post-hoc comparisons revealed that while visual search RTs for target present trials were similar for rectangle ($M = 1260$ ms) and polygon ($M = 1214$ ms) arrays ($p = 0.11$), RTs for target absent rectangle ($M = 1840$ ms) and polygon ($M = 1608$ ms) trials differed, $p < 0.0001$. However, given that twice as many shapes appeared in the rectangle search arrays relative to the polygon search arrays, it is not surprising that when the target was not present, observers were slower to terminate their search through the rectangle search arrays, although, when present, participants took a similar amount of time to locate the rectangle and polygon targets.

Composite Face Task (Test Phase)

Sensitivity (d') Data were submitted to a mixed between-subjects repeated measures 2 (training: grouped vs. ungrouped) $\times 2$ (frame: aligned vs. misaligned) $\times 2$ (congruency: congruent vs. incongruent) ANOVA, revealing a significant main effect of congruency, $F(1, 64) = 155.43$, $p \leq 0.0001$, $\eta^2_p = 0.71$, but not of training, $F(1, 64) = 0.23$, $p = 0.63$, $\eta^2_p < 0.01$, nor of frame, $F(1, 64) = 1.19$, $p = 0.28$, $\eta^2_p = 0.02$. There were no two-way interactions between any of the variables (all $p > 0.29$). However, a significant three-way interaction emerged between training, congruency, and frame, $F(1, 64) = 4.65$, $p = 0.035$, $\eta^2_p = 0.07$.

To investigate the basis of this three-way interaction, two additional two-way ANOVAs were conducted. First, a 2 (training: grouped vs. ungrouped) $\times 2$ (congruency: congruent vs. incongruent) ANOVA was conducted on the d' data from the trials with the misaligned (polygon) frames. There was a

significant main effect of congruency, $F(1, 64) = 105.27$, $p \leq 0.0001$, $\eta^2_p = 0.62$, but no main effect of training, $F(1, 64) = 0.27$, $p = 0.61$, $\eta^2_p < 0.01$. Importantly, there was a significant two-way interaction between training and congruency, $F(1, 64) = 4.07$, $p = 0.048$, $\eta^2_p = 0.06$, with participants who were trained to view the misaligned frames as a single polygon shape, showing a greater congruency effect from the task-irrelevant part than those trained to see the frames as independent rectangles. A 2 (training: grouped vs. ungrouped) $\times 2$ (congruency: congruent vs. incongruent) ANOVA on the d' data from the trials with aligned (square) frames revealed a main effect of congruency, $F(1, 64) = 107.48$, $p \leq 0.0001$, $\eta^2_p = 0.63$, but no main effect of training, $F(1, 64) = 0.08$, $p = 0.78$, $\eta^2_p < 0.01$, or interaction between training and congruency, $F(1, 64) = 0.12$, $p = 0.73$, $\eta^2_p < 0.01$ (Fig. 4).

Response Time Response time data were submitted to a mixed between-subjects repeated measures 2 (training: rectangle vs. polygon) $\times 2$ (frame: square vs. polygon) $\times 2$ (congruency: congruent vs. incongruent) ANOVA. A significant main effect of congruency emerged, $F(1, 64) = 26.75$, $p \leq 0.0001$, $\eta^2_p = 0.29$, but not of training, $F(1, 64) = 0.031$, $p = 0.86$, $\eta^2_p < 0.01$. There was a trend for a main effect of frame, $F(1, 64) = 3.35$, $p = 0.07$, $\eta^2_p = 0.17$, but no significant interactions between any of the variables (all $p > 0.30$; Table 2 of Supplementary Material section).

Discussion

Replicating findings reported in Experiment 1, holistic processing—as measured in the composite-face paradigm—was modulated by observers' experience grouping the elements of the frames within which the face parts appeared. When observers' experience fostered the perception of the differently colored, misaligned rectangles as a single multicolored polygon, faces appearing in these frames were again perceived more holistically, but in this case compared with when observers' experience fostered the perception of the framing elements as two independent rectangles.

Somewhat surprisingly, although the above patterns of findings are consistent with those from Experiment 1, unlike in Experiment 1, they were found in the sensitivity data. However, it is not unusual for findings from the composite face task to be found in response time or sensitivity data or both measures. The response time data provided very little additional insight beyond also showing a large robust effect of congruency. Thus, there was no evidence of a speed accuracy trade-off driving these findings.

Notably, the results of Experiment 2 also demonstrate that this effect of prior grouping experience was present even when the physical properties of the framing stimuli opposed

Fig. 4 Sensitivity (d') for the congruent (diamonds) and incongruent (squares) conditions, and the resulting index of holistic perception (congruency effect, filled bars, reflecting the difference between these conditions) for faces presented in aligned, same colored rectangles (left) or misaligned, different colored rectangles (right). Holistic perception of faces appearing in the misaligned, dual colored rectangles was greater among participants who had performed the pre-task encouraging perception of these framing shapes as a single polygon unit compared with those who had performed the pre-task encouraging the perception of this same frame as composed of independent rectangles. Error bars represent standard error values.

such an effect. That is, while experience encouraged the grouping of the framing elements, the image-based properties of these elements were inconsistent with such grouping; their nonuniform color and misalignment worked against the perceptual grouping of them together (Goldfarb & Treisman, 2011; Watson & Kramer, 1999).

A no-training comparison condition, not present in the current study, could have potentially provided further insight into whether the rectangle pre-task weakened holistic processing of the face by encouraging observers to view the two framing rectangles, and thus the face parts appearing in them, as separate perceptual units. Potentially noteworthy is that, in our previously published study (Curby et al., 2013), which included such a condition, the size of the congruency effect for faces appearing on the aligned/same coloured rectangles was considerably greater than that found in the rectangle pre-task condition here. This suggests that the rectangle pre-task condition served to disrupt the grouping of the face parts appearing on the aligned/same coloured framing rectangles. However, any conclusions that can be drawn by comparing these studies are limited given these experiments were conducted within different studies.

General Discussion

The results of the current experiments are consistent with two hypotheses about the nature of holistic face perception: 1) that holistic face perception is supported by—at least in part—

object-based grouping and the object-based attention this affords, whereby individual features are grouped, and attended to, as a cohesive “object” (Curby et al., 2013; Zhao, Bühlhoff, & Bühlhoff, 2015), and 2) the principles via which individual features are grouped can be modulated by experience.

Whereas the demonstration that experience can contribute to perceptual grouping and the formation of objects is not new (Kimchi & Hadad, 2002; Zemel et al., 2002), the modulation of holistic face perception by experience-dependent grouping is surprising given the often proposed impenetrability of face perception. In fact, the literature covering basic mechanisms guiding visual perception, especially with respect to the perception of objecthood, and that covering higher-level visual processes such as face perception, typically exist in parallel with little crosstalk between them.

Studies of patient groups with face perception deficits have begun to bridge this gap between the two literatures, providing intriguing support for the importance of more general-purpose object perception mechanisms in supporting characteristic features of holistic face perception. For example, patients with congenital prosopagnosia showed deficits in extracting the configural or holistic features of both face and non-face stimuli (Behrmann, Avidan, Marotta, & Kimchi, 2005). In addition, a similar study with adults with autism also showed abnormal processing of the configural features of both face and non-face stimuli, again consistent with a role of more general purpose object perception mechanisms in supporting features of holistic face perception (Behrmann et al., 2006). Thus, our findings are bolstered by studies in the broader literature that also provide evidence that more general perception mechanisms likely play an important role in supporting characteristic features of face perception.

While both Experiments 1 and 2 reported findings consistent with the impact of experience-based grouping cues on indices of holistic processing of faces, these effects were revealed in different dependent variables, namely RT and d' . Given that the literature commonly finds effects in one or the other (or both), this is not unusual. For example, the first study demonstrating that task-irrelevant face-parts interfere with part judgments of composite faces reported this in terms of a RT cost (Young et al., 1987). However, many others also have shown effects in sensitivity (see Richler & Gauthier, 2014, for a meta-analysis of studies reporting effects in d'). Notably, the two experiments reported were conducted across different institutions and thus tapped different populations of undergraduate students. The finding of converging support for the impact of experience-based grouping cues on indices of holistic processing of faces from studies conducted in different labs speaks to the robustness of this effect. Nonetheless, it remains an open question as to why observers tested in one lab primarily showed effects in RT, whereas observers tested in another lab, with an almost identical paradigm, show effects in d' . Studies targeting the underlying cause of the shifting of

effects from RT to d' are needed to better understand this paradigm.

One existing view of holistic perception is that it reflects a failure of selective attention, which results from the processing of all features of an object together, because such a strategy is necessary to extract the task-relevant information (Richler, Palmeri, & Gauthier, 2012). This strategy becomes automatized for specific categories with which people have extensive individuation experience, such as faces (Richler et al., 2012; Richler, Wong, & Gauthier, 2011). Our current findings and those from our previous study (Curby et al., 2013) suggest that the mechanisms underlying the transition from strategically motivated processing of the whole object to—following extensive experience—an automatized or obligatory holistic processing of the stimulus may be related to the strengthening of the perceptual grouping (objecthood) of the face, and the resulting operation of object-based attention.

Given the current results, it is noteworthy that commonalities between effects attributed to object-based attention and those attributed to holistic face perception can be found in the tasks used to demonstrate these effects. For example, Kramer and Jacobson (1991) used a task typically referred to as a flanker compatibility task, and the distracting flankers could be compatible or incompatible with respect to the correct response to the target stimulus. Incompatible flankers caused greater interference (relative to compatible flankers) when they were embedded in the same object as the target than when they were embedded in a different object, and this was interpreted as evidence of object-based attention. In addition, this effect was modulated by perceptual cues encouraging (common color) or discouraging (different colors) the grouping of the target and flanking stimuli. Notably, in the composite face task participants also make judgments about a target (the top or bottom half of a face) while instructed to ignore an adjacent distractor (the other, task-irrelevant, half of the face). As in Kramer & Jacobson's (1991) classic task, the task-irrelevant (distractor) face half could be compatible or incompatible with respect to the correct response to the task-relevant (target) face half. A key manipulation of the standard composite task is that the distracting face half is either aligned, binding with the target face half to make a coherent face, or misaligned, disrupting the perception of the face as a single perceptual unit. Thus, manipulations that discourage or encourage the perceptual grouping of face parts may modulate indices of holistic processing via their impact on the employment of object-based attention. Specifically, such manipulations likely facilitate or impair participants' ability to selectively attend to the target face part, not unlike what is observed in paradigms that demonstrate object-based attentional selection.

A further, striking similarity between perceptual grouping and holistic face perception is the time-course under which

these processes operate. Feldman (2007) mapped the time-course of object formation, finding that representations of fully bound visual objects, i.e., "objecthood," emerge after approximately 200 ms of processing. This proposed time-course maps onto the time-course of holistic face perception (Richler, Mack, Palmeri, & Gauthier, 2011), as well as onto the timing with which manipulated perceptual grouping cues impact holistic perception (Curby et al., 2013). Furthermore, Kimchi & Hadad (2002) demonstrated that the effects of past experience and those of physical, image-based cues were evident on the same time scale. They suggest that perceptual organization, like that which occurs with the grouping of elements/features into objects, is best characterized as an interactive process between physical, image-based stimulus properties and experience-based stimulus representations.

Recent studies also have highlighted a potential role of attention in establishing holistic processing of faces. Specifically, the impact of learned attentional strategies on holistic perception was assessed using a modified composite task with novel race face stimuli (Chua, Richler, & Gauthier, 2014). Participants in this study individuated exemplars of these novel race stimuli where either the top or bottom face halves contained diagnostic (distinguishing) information. Holistic processing was measured using the composite face task with stimuli made from either diagnostic or nondiagnostic parts from the two stimulus sets. In this study, holistic processing was only found for the stimuli made from diagnostic parts, even though observers have never before seen these parts as part of the same stimulus. This finding was replicated in a follow-up study with non-face-like novel stimuli, Greebles (Chua, Richler, & Gauthier, 2015). These findings suggest that learned contingencies that shape the way faces are attended may play a central role in establishing holistic perception for face and non-face object of expertise.

Studies also have revealed that learned contingencies about the likely location of task-relevant information modulate object-based attention effects (Shomstein & Behrmann, 2008). Similar to the findings regarding holistic processing of the novel race or Greeble stimuli, when learned contingencies suggested that task-relevant information was not likely to be found within the same object, the same-object attentional advantage (a classic marker of object-based attention) was reduced. This finding led Shomstein and Behrmann (2008) to propose that object-based attentional advantages are the result of attentional prioritization. This prioritization is guided both by the basic perception properties of the stimulus (strength of objecthood) and by learned probabilistic contingencies.

While we focus here on the potential impact and role of experience in shaping object-based attention to faces and thus holistic processing, it is equally important to note the importance of stimulus-level features and cues in shaping objecthood and object-based attention. Consistent with the relevance of such factors in understanding holistic perception,

a recent paper demonstrated the importance of gestalt grouping cues in driving holistic perception (Zhao, Bühlhoff, & Bühlhoff, 2015). Specifically, this study found evidence of holistic perception of non-face patterns that were strong in gestalt grouping cues. This study, together with the study reported here, highlight the potential importance of more general-purpose perceptual processing mechanisms in driving and supporting holistic face perception.

Of note is how the perceptual grouping of the face parts, and thus object-based attention to the face, was manipulated indirectly by manipulating the perception of the frame within which the face parts appeared. This was a necessary design feature. Numerous previous studies have introduced manipulations that disrupt the objecthood of the face and the experience-based grouping of the face, such as physically misaligning the face parts (disrupting good continuation of form), putting the face parts on different depth planes (Taubert & Alais, 2009), increasing the spacing between top and bottom face halves (Taubert & Alais, 2009), or rotating the face 180 degrees (experience-based effects on grouping are orientation-specific; Vecera & Farah, 1997). However, in such cases, the reduction in the holistic perception of faces cannot be conclusively attributed to the disruption of the objecthood of the face because in these cases the physical configuration of the face has also been disrupted. This is particularly problematic as the configuration of the face plays a key role in many prominent theories of face perception. For example, norm-based coding theories of face perception also predict a disruption to face perception if the physical configuration of the face is disrupted, such as in the case of physically misaligning the top and bottom parts of composite faces (Rhodes & Jeffery, 2006). However, this and other existing theories of face perception (Farah, Wilson, Drain, & Tanaka, 1998; Tanaka & Farah, 1993) would have difficulty accounting for manipulations that disrupt holistic face perception while the configuration of the face remains intact and unchanged, as in the current study.

Depending on how the observer attends to the stimuli, it is possible that they perceive the face parts and the rectangles that frame them as separate objects or as part of the same object. The presence of the expected effect of the framing rectangles in shaping objecthood (and thus object-based attention) suggests that the frame and face elements are interacting, at least at some level. One might suggest that the frame is still perceived independently of the face parts and that the grouping of the framing shapes only impacts the perceived grouping of the face parts because they are overlapping in space. However, object-based attention studies that demonstrate that observers can independently select one of two superimposed objects (e.g., O'Craven, Downing, & Kanwisher, 1999; Serences, Schwarzbach, Courtney, Golay, X., & Yantis, 2004) suggests that it is more likely that the effect arises via

integration, at least to some degree, of the face parts and the rectangle frames.

On a related note, it is possible that the mechanisms supporting experience-based perceptual grouping interact with, while operating in parallel with, other (face-specific) mechanisms supporting holistic face perception. This possibility would be insightful in terms of our understanding of holistic face perception: if these mechanisms work together with other, more face-specific, mechanisms to produce holistic effects, this would suggest that holistic face perception is penetrable by basic perceptual mechanisms. Given prominent current models of face perception have suggested that holistic face perception is supported by an encapsulated, domain-specific module that is immune to the influence of non-face processing mechanisms, this possibility would also result in new insights, extending our understanding of holistic face perception. Regardless, further work is necessary to determine the precise role and degree to which experience-based strengthening of the objecthood of the face contributes to the characteristic holistic nature of face perception.

In summary, object-based attention shares a number of characteristics with holistic perception. For example, studies of object-based attention demonstrate that the distribution of attention throughout an object can be shaped by the probability of the location of task-relevant information (Shomstein & Yantis, 2002) and our experience with the object (Vecera & Farah, 1997), similar to recent findings regarding holistic perception effects (Chua et al., 2015). Furthermore, object-based attention has also been shown to operate in an automatic, inflexible fashion even when such an attentional distribution is non-optimal and interferes with task performance (Kramer & Jacobson, 1991), not unlike what is observed when individuals are asked to make judgments about only part of a face in the composite face paradigm (Gauthier et al., 2003). Thus, object-based attention shares many of the characteristics necessary for any candidate mechanism contributing to holistic perception effects.

In conclusion, the findings reported here not only provide insight into the holistic perception of faces but also into mechanisms related to the perception of objecthood and the resulting object-based attention. Whereas much is known about the role of image-based properties on the perceptual grouping of features into objects, and also that such mechanisms are influenced by experience, the full potential of experience to impact the perception of objecthood, especially in extreme cases such as face and non-face expert perception, is relatively unknown. These findings support the possibility that holistic perception is, at least in part, a consequence of extensive experience perceptually grouping the features within faces, resulting in the obligatory selection of the face as an attentional unit.

Acknowledgments This research was supported by a grant from the Australian Research Council (ARC) (DE130100969). We would also like to thank Denise Moerel for help with data collection.

References

- Amishav, R., & Kimchi, R. (2010). Perceptual integrality of componential and configural information in faces. *Psychonomic Bulletin & Review*, 17(5), 743–748.
- Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., & Minshew, N. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia*, 44(1), 110–129.
- Behrmann, M., Avidan, G., Marotta, J. J., & Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *Journal of Cognitive Neuroscience*, 17(7), 1130–1149.
- Boggan, A. L., Bartlett, J. C., & Krawczyk, D. C. (2012). Chess masters show a hallmark of face processing with chess. *Journal of Experimental Psychology: General*, 141(1), 37–42.
- Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition*, 38(2), 109–144.
- Bukach, C. M., Cottle, J., Ubiwa, J., & Miller, J. (2012). Individuation experience predicts other-race effects in holistic processing for both Caucasian and Black participants. *Cognition*, 123(2), 319–324.
- Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, 45(4), 431–448.
- Calder, A., Rhodes, G., Johnson, M., & Haxby, J. (Eds.). (2011). *Oxford Handbook of Face Perception*: Oxford University Press.
- Chua, K. W., Richler, J. J., & Gauthier, I. (2014). Becoming a Lunari or Taiyo expert: Learned attention to parts drives holistic processing of faces. *Journal of Experimental Psychology: Human Perception & Performance*, 40(3), 1174–1182.
- Chua, K. W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention to parts. *Journal of Experimental Psychology: General*, 144(4), 723–729.
- Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception & Performance*, 35(1), 94–107.
- Curby, K. M., Goldstein, R. R., & Blacker, K. (2013). Disrupting perceptual grouping of face parts impairs holistic face processing. *Attention, Perception, & Psychophysics*, 75(1), 83–91.
- Curby, K. M., Johnson, K. J., & Tyson, A. (2012). Face to face with emotion: Holistic face processing is modulated by emotional state. *Cognition & Emotion*, 26(1), 93–102.
- Egley, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123(2), 161–177.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, 105(3), 482–498.
- Feldman, J. (2007). Formation of visual "objects" in the early computation of spatial relations. *Perception & Psychophysics*, 69(5), 816–827.
- Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, 6(4), 428–432.
- Goldfarb, L., & Treisman, A. (2011). Does a color difference between parts impair the perception of a whole? A similarity between simultanagnosia patients and healthy observers. *Psychon Bulletin & Review*, 18(5), 877–882.
- Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, 13(5), 505–512.
- Hollingworth, A., Maxcey-Richard, A. M., & Vecera, S. P. (2012). The spatial distribution of attention within and across objects. *Journal of Experimental Psychology: Human Perception & Performance*, 38(1), 135–151.
- Kimchi, R., & Hadad, B.-S. (2002). Influence of past experience on perceptual grouping. *Psychological Science*, 13(1), 41–47.
- Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: the role of objects and proximity in visual processing. *Perception & Psychophysics*, 50(3), 267–284.
- Matsukura, M., & Vecera, S. P. (2006). The return of object-based attention: Selection of multiple-region objects. *Perception & Psychophysics*, 68(7), 1163–1175.
- Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology*, 19, 545–567.
- Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition*, 14(1), 55–73.
- Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, 17(7), 608–615.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature* 401, 584–587.
- Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Science*, 9, 441–474.
- Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, 46(18), 2977–2987.
- Richard, A. M., Lee, H., & Vecera, S. P. (2008). Attentional spreading in object-based attention. *Journal of Experimental Psychology: Human Perception & Performance*, 34(4), 842–853.
- Richler, J. J., Bukach, C. M., & Gauthier, I. (2009). Context influences holistic processing of nonface objects in the composite task. *Attention, Perception, & Psychophysics*, 71(3), 530–540.
- Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, 140(5), 1281–1302.
- Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research*, 51(3), 333–342.
- Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, 3, 553.
- 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, & Cognition*, 34(6), 1356–1368.
- Richler, J. J., Wong, Y. K., & Gauthier, I. (2011). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science*, 20(2), 129–134.
- Serences, J. T., Schwarzbach, J., Courtney, S. M., Golay, X., & Yantis, S. (2004). Control of object-based attention in human cortex. *Cerebral Cortex* 14(12), 1346–1357.
- Shomstein, S., & Behrmann, M. (2008). Object-based attention: Strength of object representation and attentional guidance. *Perception & Psychophysics*, 70(1), 132–144.
- Shomstein, S., & Yantis, S. (2002). Object-based attention: Sensory modulation or priority setting? *Perception & Psychophysics*, 64(1), 41–51.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A*, 46(2), 225–245.
- Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, 93(1), B1–B9.

- Taubert, J., & Alais, D. (2009). The composite illusion requires composite face stimuli to be biologically plausible. *Vision Research*, 49(14), 1877–1885.
- Vecera, S. P., & Farah, M. J. (1997). Is visual image segmentation a bottom-up or an interactive process? *Perception & Psychophysics*, 59(8), 1280–1296.
- Vickery, T. J., & Jiang, Y. V. (2009). Associative grouping: Perceptual grouping of shapes by association. *Attention, Perception, & Psychophysics*, 71(4), 896–909.
- Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics*, 61(1), 31–49.
- Young, A. W., Hellawell, D., & Hay, D. (1987). Configural information in face perception. *Perception*, 10, 747–759.
- Zemel, R. S., Mozer, M. C., Behrmann, M., & Bavelier, D. (2002). Experience-dependent perceptual grouping and object-based attention. *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), 202–217.
- Zhao, M., Bülthoff, H. H., & Bülthoff, I. (2015). Beyond faces and expertise: Facelike holistic processing for nonface objects in the absence of expertise. *Psychological Science*. doi:[10.1177/0956797615617799](https://doi.org/10.1177/0956797615617799)
- Zhao, M., Hayward, W. G., & Bulthoff, I. (2014). Holistic processing, contact, and the other-race effect in face recognition. *Vision Research*, 105, 61–69.