ARTICLE IN PRESS

Consciousness and Cognition xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Perceptual averaging of facial expressions requires visual awareness and attention

Elric Elias*, Lauren Padama, Timothy D. Sweeny

Department of Psychology, University of Denver, United States

ARTICLE INFO

Keywords:
Face perception
Awareness, object substitution masking
Consciousness
Attention

ABSTRACT

Humans, as highly social animals, are regularly exposed to the faces of conspecifics—often more than one at a time. This feature of social living is important for understanding face perception, not just because it means that information from faces is available in bulk, but also because it changes the way individuals are perceived. For instance, when two faces are seen nearby one another, they tend to look like each other. This phenomenon of perceptual averaging is robust when both faces are seen and attended. But in everyday life, some faces may not receive the full benefit of attention, or they may not be visible at all. We evaluated whether perceptual averaging of relatively complex and simple information on faces, including facial expression and head orientation, can still occur even in these circumstances. In particular, we used object-substitution masking (OSM) and a dual-task designed to disrupt visual awareness and attention, respectively, during evaluations of briefly presented face pairs. Disruptions of awareness or attention eliminated averaging of facial expression, whereas orientation averaging persisted in spite of these challenges. These results demonstrate boundary conditions for the process of perceptual averaging. More generally, they provide insight into how the visual system processes multitudes of objects, both simple and complex, both with and without attention and awareness.

1. Introduction

People are sometimes seen one at a time. A friend's smile, for example, is easy to recognize across the room or in a video chat. In situations like these, sophisticated neural mechanisms support rapid analysis of a face's structure (Young & Yamane, 1992), identity (Hasselmo, Rolls, & Baylis, 1989), and emotion (Streit et al., 1999), sometimes even if that face's expression is not consciously seen (Adams, Gray, Garner, & Graf, 2010). Much of the psychological and vision research that focuses on the process of face perception is constrained to this solitary level of analysis. Yet humans are social creatures and people are often seen in the presence of others (Van Vugt & Kameda, 2012). Understanding how faces are encoded and then seen in these more complex circumstances is not simply a matter of extrapolating knowledge about how faces are perceived one at a time. Rather, the way people are seen depends on additional neural computations that take into account the presence of others nearby (Cosmides & Tooby, 2005). Work that highlights person-perception as it commonly occurs—in the context of other people—is crucial for a complete understanding of how faces are represented and then perceived by the visual system.

Toward this end, a great deal of attention has been focused on how visual information transcends individuals to the level of the collective (e.g., Alvarez, 2011; Whitney, Haberman, & Sweeny, 2014). For example, the process of ensemble coding allows an observer to appreciate the average emotion of an entire crowd at the expense of knowledge about its constituents (e.g., Elias, Dyer &

* Corresponding author at: University of Denver, Department of Psychology, 2155 S Race St, Frontier Hall, Denver, CO 80210, United States.
E-mail address: elric.elias@du.edu (E. Elias).

https://doi.org/10.1016/j.concog.2018.03.005
Received 25 September 2017; Received in revised form 27 December 2017; Accepted 13 March 2018

Sweeny, 2017; Haberman & Whitney, 2007, 2009). Less attention has been focused on a separate but possibly related process—*perceptual averaging*—in which visual information is transferred *between* individuals, within the context of a crowd (Sweeny, Grabowecky, Paller, & Suzuki, 2009; Sweeny, Grabowecky, Kim, & Suzuki, 2011). In perceptual averaging, two spatially distinct faces can be made to appear similar to each other. For example, the anger of one face may appear to spread to a neutral face seen nearby. Although some details about the perceptual averaging of facial expression are clear—for example, it is likely to be, at least in part, a consequence of the large receptive fields of high-level ventral visual neurons—many questions remain about its mechanisms, its limitations, and when it can be expected to occur. The goal of the current investigation was to begin to bridge this gap by evaluating how perceptual averaging between multiple faces may depend on visual awareness and attention. In doing so, we hoped to better characterize the automaticity of face-specific perceptual averaging, determining whether it operates quickly, potentially based on feedforward visual representation characteristic of unconscious vision, or if it instead requires more in-depth and iterative representation associated with visual awareness (Pascual-Leone & Walsh, 2001; Silvanto, Cowey, Lavie, & Walsh, 2005; Juan & Walsh, 2003). Before discussing how perceptual averaging of facial expressions may depend on visual awareness, we first describe the phenomenon more generally from both neurophysiological and psychophysical perspectives.

Although the encoding and recognition of facial expressions of emotion is a distributed process (e.g., Pallett & Meng, 2013; Streit et al., 1999), a key stage of this analysis occurs in temporal visual cortex (e.g., Hasselmo et al., 1989). Neurons in high-level areas of the ventral visual stream like inferotemporal cortex (IT) are known to have very large receptive fields (e.g., Boussaoud, Desimone, & Ungerleider, 1991; Chelazzi, Duncan, Miller, & Desimone, 1998; Desimone & Gross, 1979; Niemeier, Goltz, Kuchinad, Tweed, & Vilis, 2004; Op De Beeck & Vogles, 2000). These large receptive fields, which often take up the entire contralateral visual hemifield, provide a notable computational benefit—they allow cells to respond consistently to a complex pattern, like a single face, despite dramatic changes to its location. However, this position-invariant coding introduces a cost of poor spatial resolution, which can be problematic when multiple objects or faces appear within a neuron's receptive field. Overcoming this tradeoff is often as simple as engaging selective attention, which can suppress the representation of the unattended object and allow such a neuron to respond as if only one object were in its receptive field (Chelazzi et al., 1998). But of course, selective attention cannot always be engaged because the locations of important objects are not always certain, attention must sometimes be distributed, and perception can often be fleeting. In these kinds of circumstances, ventral visual neurons are unable to resolve the multiple objects within their receptive fields, and they respond as if they "average" across those patterns (e.g., Kastner et al., 2001; Miller, Gochin, & Gross, 1993; Rolls & Tovee, 1995; Sato, 1989; Zoccolan, Cox, & Di Carlo, 2005). For example, when a pattern that elicits a strong response from a particular neuron (at least when seen in isolation, e.g., 30 spikes/s) is seen nearby a pattern that does little to push the neuron past its spontaneous firing rate (e.g., 10 spikes/s), this unit will produce an intermediate response (e.g., 20 spikes/s). This neural averaging does not occur when selective attention is engaged on one object from a pair, nor does it occur when the two objects appear in different receptive fields of ventral visual neurons (i.e., one on either side of the vertical meridian). Crucially for the current investigation, this process of neural averaging has a striking perceptual consequence.

When a face with a surprised expression is seen nearby a second face with a happy (or angry) expression, for example, the affect from each face appears to spread to the other (Sweeny et al., 2009, 2011). In this case, the happy face would appear less happy next to a surprised face, and vice versa. Crucially, this perceptual averaging occurs most strongly when a pair of faces is seen within a visual hemifield compared to when the faces appear in separate visual hemifields, illustrating that perceptual averaging can be constrained by the spatial selectivity (or lack thereof) of high-level visual neurons. Furthermore, the fact that averaging, both neural and perceptual, occurs when selective attention is *not* engaged on any individual face suggests that this process may operate as if by default, a natural consequence of the visual system's architecture that occurs any time information about multiple objects is present yet unbound by selection. Here, we put this notion of default averaging to a more stringent test, examining whether or not it would occur even when one face from a pair is not visible. This is important to examine, since lapses in visual awareness can be reasonably expected to occur in cluttered and rapidly changing visual scenes where multiple people are so often encountered. Moreover, faces, especially those with emotional expressions, are known to engage rapid but coarse visual processing even when they are not visible (for reviews, see Axelrod, Bar, & Rees, 2015; Faivre, Berthet, & Koudier, 2012). Many masking techniques could have allowed us to manipulate visual awareness in our investigation, but one in particular was best suited for this investigation—object-substitution masking (OSM).

Object-substitution masking (OSM) is a unique and powerful tool for disrupting perception (Enns, 2004; Enns & Di Lollo, 1997). In most OSM paradigms, a target object is shown for a brief amount of time, typically flanked by four masking dots. When these dots persist after the target disappears, even for just a fraction of a second, they can disrupt discrimination of that object's features and even eliminate awareness of its presence (although these effects can be independent; Gellatly, Pilling, Cole, & Skarratt, 2006). Although recent evidence suggests that OSM is not special in terms of its interaction with spatial attention (Argyropoulos, Gellatly, Pilling, & Carter, 2013; Filmer, Mattingley, & Dux, 2014, 2015; Goodhew & Edwards, 2016; Pilling, Gellatly, Argyropoulos, & Skarratt, 2014), it can still be differentiated from other types of masking in terms of its time course (Enns, 2004). More importantly, OSM is known to have a relatively late stage of interruption (Chakravarthi & Cavanagh, 2009), and relative to other types of visual masking (e.g., sandwich masking; Harris, Wu & Woldorff, 2011, backward masking; Woodman & Luck, 2003), we expected OSM to be the least likely to interfere with visual processing of facial expression over and above disrupting visual awareness. Below, we provide some background on the potential mechanisms of OSM in order to place our design and hypotheses in a deeper theoretical context.

Both early and more recent accounts of OSM (e.g., Di Lollo, 2014; Enns & Di Lollo, 1997; Enns, 2004) highlight the way it appears to selectively disrupt the kind of re-entrant communication between higher- and lower-level visual areas (e.g., extrastriate areas and V1) that appears to be necessary for visual awareness of objects (Lamme, Super, & Spekreijse, 1998; Pascual-Leon & Walsh, 2001; Silvanto, Cowey, Lavie, & Walsh, 2005). More specifically, these accounts of OSM propose that when information about a masked object (e.g., a face surrounded by masking dots) arrives in higher-level stages of vision via a feedforward sweep of processing, the visual system generates an initial hypothesis about that object's appearance and location. This hypothesis is then tested by sending information about that object (in this case, the face *and* masking dots) back to earlier visual areas for comparison with the most current visual input or, at worst, a decaying trace of that object from a moment ago. According to this account, when this feedback matches the current input, a phenomenological experience of that object is likely to follow shortly thereafter. OSM may eliminate awareness by disrupting this matching process. For example, if only the masking dots are visible when feedback activity arrives, the reentrant representation of the face and dots would not match the current input. In this case, the visual system would abandon its more complex hypothesis of having encountered an object (in this case, a face), and instead allow visual awareness of the masking dots alone (Di Lollo, Enns, & Rensink, 2000). Several investigations provide support for this iterative account (Boehler, Schoenfeld, Heinze, & Hopf, 2008; Pascual-Leone & Walsh, 2001), which also seems to be important during face perception (De Haan, Pascalis, & Johnson, 2002). Crucially, this substitution model of OSM features little or even zero interference both at the level of local contour interactions (Di Lollo et al., 2000) and across the initial wave of feedforward activity from V1 to higher-level visual areas (Enns, 2004; Goodhew, Dux, Lipp, & Visser, 2012; Jannati, Spalek, & Di Lollo, 2013; Kotsoni, Csibra, Mareschal, & Johnson, 2007).

A second and more recent account of OSM, often referred to as the object-updating account, is quite similar to the substitution account described above, except that rather than substituting one representation for another, a single, ongoing representation of the target object is updated to only include the masking dots (Goodhew, 2017; Goodhew, Edwards, Boal, & Bell, 2015; Lleras & Moore, 2003; Moore & Lleras, 2005). Iterative processing and minimal interference are still plausible mechanisms within this updating account (Filmer et al., 2015; Pilling et al., 2014).

A third characterization of OSM proposes that local interference, like lateral inhibition or the addition of noise from the masking dots, may degrade processing and perception of the target object (Bridgeman, 2006; Macknik & Martinez-Conde, 2007; Poder, 2012). This account tends not to include iterative processing as a central piece of the puzzle, and instead focuses on the accumulation of interference in object representation during feedforward processing.

Importantly, we note that all three of these accounts (1) allow that a feedforward sweep of analysis should occur during OSM (even if that sweep is somewhat disrupted), and (2) predict that when OSM eliminates awareness of an object, that masked object's ability to nevertheless influence increasingly complex perceptual judgments should depend, at least to some extent, on the strength of its representation as it moves up this feedforward sweep. In the case of OSM, disruption of this feedforward sweep should be limited (or at least no worse) relative to other types of masking. In other words, OSM allowed us to manipulate awareness of a face while, at the very least, minimizing interference to its initial visual representation, thus increasing the likelihood that we would be able to capture an effect of face-specific perceptual averaging without visual awareness, if such a phenomenon were possible.

There are a few reasons to expect that perceptual averaging between a pair of faces might still occur even when one face is not visible. First, complex categorizations that require processing in high-levels of the ventral visual pathway (e.g., detecting animals or faces in natural visual scenes) can occur exceptionally quickly, presumably based on the kind of feedforward processing that may be preserved in OSM (VanRullen, & Thorpe, 2001; VanRullen & Koch, 2003; Rousselet, Macé, & Fabre-Thorpe, 2003). Second, population codes in macaque STS can discriminate complex images, including faces, at surprisingly rapid timescales consistent with feedforward processing (Keysers, Xiao, Foldiak, & Perrett, 2001). Third, even when OSM eliminates visual awareness of an object, relatively simple information about its features, like size or orientation, can still bias the perception of other objects seen clearly and nearby (Choo & Franconeri, 2010; Jacoby, Kamke & Mattingley, 2013). Fourth, integration of information across faces seems to operate even when observers are unable to discriminate information about individual faces during ensemble coding (e.g., Haberman & Whitney, 2011), and when a face's appearance is degraded (but still detected) during crowding (Fischer & Whitney, 2011).

Nevertheless, there are also reasons to be skeptical that perceptual averaging could operate without visual awareness of a face. In a recent investigation, we found that although interactions between complex shapes processed in intermediate stages of vision (e.g., V3/VP & V4, Dumoulin & Hess, 2007) persisted even when one shape was rendered invisible by OSM, these effects were notably weakened compared to when OSM was absent (Sweeny, D'Abreu, Elias, & Padama, 2017). Additionally, OSM has been shown to eliminate representation of complex objects, at least in terms of activity measurable by fMRI (Carlson, Rauschenberger, & Verstraten, 2007), as well as differences in the face-specific N170 component during discrimination of faces and houses (Reiss & Hoffman, 2007). This dovetails with findings that, during metacontrast masking, activity in the fusiform gyrus (an area important for face perception), but not area V4 (an area important for shape perception) seems to be related to visibility (Haynes, Driver, & Rees, 2005). Neural noise is known to accumulate as information travels up the visual hierarchy (Faisal, Selen, & Wolpert, 2008). It may be the case that, compared to simpler visual features, decaying feedforward information about faces has insufficient potency to drive the process of perceptual averaging by the time it reaches high-levels of the ventral visual pathway. Although these points converge to suggest that expression averaging should not occur during OSM, we did not consider this outcome to be a forgone conclusion, particularly because processing of emotional content is known to occur independently of determining that a face is present (Hung et al., 2010), and neuroimaging may sometimes fail to capture lingering activity capable of influencing perception (Carlson et al., 2007).

In the current investigation, we presented pairs of faces, one with a happy or angry expression and one with a neutral expression. We predicted that, as in Sweeny et al. (2009), the emotion on the faces would appear to spread more readily when both faces appeared within a visual hemifield than when they appeared in separate visual hemifields. In Experiment 1, we predicted that this effect of perceptual averaging would be severely weakened or even eliminated when OSM disrupted awareness of one face in the pair. Since OSM can independently interfere with awareness and visual processing (Gellatly et al., 2006), we carefully examined the strength of perceptual averaging taking into account each observer's reports of visual awareness on a trial-by-trial basis, an approach that has been used surprisingly infrequently in studies of OSM (e.g., Harris, Ku, & Woldorff, 2013; Harrison, Rajsic, & Wilson, 2016; Gellatly et al., 2006; Kahan & Enns, 2010; Prime, Pluchino, Eimer, Dell'Acqua, & Jolicquer, 2011). In Experiment 2, we examined the extent that perceptual averaging of emotion could occur when both faces were visible, but attention was occupied by an additional orientation-judgment task. Overall, this design allowed us to accomplish our main objective—determining the extent to which perceptual averaging of facial expression can occur during disruptions of visual awareness and attention—while at the same time providing new insight into the extent to which visual processing persists during OSM.

2. Experiment 1

2.1. Material and methods

2.1.1. Observers

Thirty undergraduate students with normal or corrected-to-normal vision from the University of Denver participated in Experiment 1. A previous investigation with a similar design and analysis, but no masking condition, had sufficient power to detect an averaging effect with twenty-four observers completing just 56 trials (Sweeny, Grabowecky, Paller, & Suzuki, 2009). We were uncertain about the strength of potential effects in the masking condition and how often masking would be successful in the current investigation. We thus collected data from 30 observers, each completing 640 trials.

2.1.2. Stimuli

Our stimulus set included color photographs of four Caucasian male actors from the NimStim face set (Tottenham, Borscheid, Ellertsen, Marcus, & Nelson, 2009). We selected neutral, happy, and angry closed-mouth expressions from each of these four actors for a total of 12 face images. We made adjustments to these original faces for a previous investigation (Elias et al., 2017), replacing the background of each image with uniform gray (R/B/G = 170/170/170; luminance = 43.65 cd/m^2) and smoothing the contours around each actor's head to reduce aliasing. A norming experiment from this previous investigation also confirmed that, when shown for 500-ms at fixation, these emotional exemplars were indeed perceived to be more intense than the neutral images (see Elias et al., 2017, for details). We were thus confident that these stimuli would be appropriate for measuring effects of emotional averaging in the current investigation.

A pair of faces appeared on every trial. Faces in the pair appeared at two of four peripheral locations along an invisible iso-acuity ellipse. Placement along this ellipse ensured that faces would be perceived with comparable acuity (Rovamo & Virsu, 1979). Face pairs were arranged either vertically, with both to the left or to the right of a central fixation point (the *within-hemifield* condition) or horizontally, with one face to the left of fixation and one to the right, with both above or below fixation (the *between-hemifield* condition) (Fig. 1A). All faces subtended a visual angle of 4.88° × 5.5°. The center of each face was 4.6° from fixation. The distance

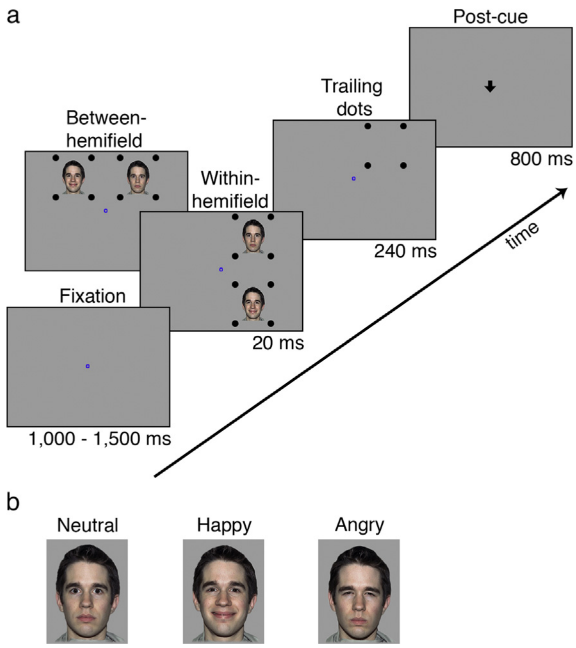

Fig. 1. (A) A typical trial sequence. Each trial contained a pair of faces. One face always displayed a neutral expression and the other face displayed either a happy or an angry expression. In the within-hemifield condition, the pair of faces appeared entirely in the left or right visual hemifield. In the between-hemifield condition, one face appeared in the left visual field and one face appeared in the right visual field, both above or both below fixation. Each face was surrounded by a quartet of black masking dots. On some trials, a set of masking dots remained on the screen after the offset of the faces, presumably masking the face that appeared in that location—the *masking* condition. On other trials, all masking dots offset with the faces and were followed by a blank screen instead of trailing dots—the *no-masking* condition. The post-cue, a centrally presented arrow, indicated which face from the pair the observer should rate in terms of its emotional valence. (B) Examples of one identity from the face set displaying neutral, happy, and angry expressions.

between face pairs was closer along the horizontal axis (6.4°, center to center) than along the vertical axis (7.2°, center to center). Notably, these inter-stimulus distances should not have been sufficient to produce crowding (Levi, 2014; Petrov & Meleshkevich, 2011). We used vertical and horizontal spatial arrangements so that we could separately measure the strength of perceptual averaging when a pair of faces fell within versus between visual hemifields, and thus either within or across the receptive field(s) of populations of ventral visual neurons, respectively.

Separating faces in the within-hemifield condition by a greater spatial distance than faces in the between-hemifield condition served two important purposes. First, it allowed us to re-affirm that, as in Sweeny et al. (2009), the strength of perceptual averaging is determined primarily by large receptive fields of ventral visual neurons, and not spatial proximity, per se. In this case, averaging should be stronger in the within-hemifield condition. Alternatively, if the strength of perceptual averaging were simply determined by spatial proximity, it would be stronger in the between-hemifield condition in which faces were physically closer but separated by a greater distance in terms of cortical representation (with each face represented in a different cortical hemisphere). Second, the between-hemifield trials (and the close proximity of the faces in them) served as a sort of control condition, allowing us to confirm that data indicative of perceptual averaging did not emerge from response biases or accidental key presses on trials in which both faces from a pair were visible. Accidentally reporting the valence of the un-cued face, even on a fraction of trials, would produce results suggestive of averaging, regardless of which face was cued. Critically, if this kind of non-perceptual factor were to influence our results, it would be equally likely to occur in both spatial arrangements, if not more strongly on between-hemifield trials in which the shorter distance between the faces could reasonably be expected to lead to more confusion and response errors. The particular pattern of results we predicted—stronger perceptual averaging on within-hemifield trials when both faces were visible—would only occur as the consequence of neural averaging. However, obtaining this clear measure of perceptual averaging in the absence of OSM (henceforth referred to as *no-masking trials*) was not our main objective. It simply allowed us replicate our previous findings and measure the extent to which perceptual averaging occurred in the context of our new design and with our stimuli, when masking did not occur. With this baseline measure of averaging in hand, we could then evaluate the extent to which OSM and disruptions of visual awareness influenced perceptual averaging, if at all.

We did not include a condition in which only a single face was presented. Although doing so would have allowed us to obtain a baseline measure of how expressions were perceived in isolation, and how masking degraded this perception, we were primarily interested in the interactive effect between pairs of faces. Plus, a single-face condition would have been difficult to compare with our two-face conditions due to different demands of spatial attention and the relevance of the post-cue (see below). We thus focused on interactions between pairs of faces throughout the investigation.

Experiments were conducted on a CRT monitor with a refresh rate of 100 Hz at a viewing distance of 57 cm. Experiments were run using Matlab using the Psychophysics Toolbox (Brainard, 1997).

2.1.3. Procedure

The experimenter provided detailed instructions and allowed each observer to participate in an unlimited number of practice trials until he or she felt ready to start the experiment. Observers were told to maintain their gaze on the fixation point at the center of the screen, but also to let their attention spread across the entire display.

At the beginning of each trial, a central fixation point (0.2° × 0.3°) appeared for a duration between 1000- and 1500-ms, which varied randomly across trials. Next, a pair of faces appeared for 20-ms in either a within-hemifield or between-hemifield spatial arrangement. Spatial arrangement was randomly determined on each trial. One face in each pair always portrayed a neutral expression and the other face portrayed either a happy or angry expression. The identity of each face was randomly selected on every trial, though a single trial never featured two faces with the same identity. Both faces in a pair were surrounded by four black dots, one just outside each corner of the face, each equidistant (2.5°) from that face's center (Fig. 1A). The dots onset with the faces, which made it impossible for observers to determine which face from the pair would be masked (if at all) while they were still on the screen. On *no-masking trials*, all dots offset simultaneously with both faces. On *masking trials*, the dots surrounding one face in the pair had a delayed offset, remaining visible for 240-ms beyond the offset of the face they surrounded (Fig. 1A). Previous investigations have shown that feedback activity tends to arrive in early visual areas with a latency of 80–120-ms (Jannati et al., 2013), and that the timing of this reentrant activity is related to the effectiveness of OSM (Kotsoni et al., 2007). Thus, according to re-entrant accounts of OSM, our 240-ms lag time should have been more than adequate to induce masking. The location of the trailing mask was counterbalanced so that it appeared around each face in each spatial arrangement an equal number of times.

Immediately after the offset of the stimuli (the faces and masks, or the trailing mask alone), a black arrow cue (1.2° × 1.2°, luminance = 0.86 cd/m^2) appeared at the center of the screen for 800 ms. The arrow pointed either to the left or right on between-hemifield trials or up or down on within-hemifield trials, indicating (with equal probability) the neutral or emotional face as the face to be rated. On masking trials, only the unmasked face was cued. Note that it would have been impossible for observers to predict the location of the cued face when each pair was on the screen. This was true even on masking trials, since the location of the masked face was not revealed until both faces had disappeared. This uncertainty, in combination with the instructions to focus attention on the central fixation point and the unpredictability of the spatial arrangements, made systematic deployment of focal attention to the cued face impossible.

A response screen then appeared immediately after the offset of the arrow. One rating scale appeared near the top of the screen instructing observers to rate the valence and intensity of the cued face (*1 = most positive, 2 = moderately positive, 3 = moderately negative, 4 = most negative,* or *1 = most negative, 2 = moderately negative, 3 = moderately positive, 4 = most positive*, counterbalanced across observers). A second rating scale appeared simultaneously near the bottom of the screen instructing observers to report on their visual awareness on each trial. The text "How many faces did you see clearly?" appeared, along with the text "1 (left arrow), 2 (right arrow).” Observers first rated the emotional intensity of the cued face by pressing one of four keys on the keyboard. Observers were permitted to revise this response as many times as necessary, in case their initial response was provided in error. Then, observers indicated how many faces were clearly visible on that trial (1 or 2) by pressing the left or right arrow, respectively. This response ended the trial. We note that our measure of awareness was subjective, and thus it is unclear how the masking we used would have influenced awareness in a more objective test of face-detection.

Observers had unlimited time to respond. All 30 observers completed 640 trials, which were counterbalanced across the within- and between-hemifield spatial arrangements. The emotional expression (happy or angry) paired with the neutral face, the location of the emotional face (top, bottom, left, or right), and the cued face (neutral or emotional) were counterbalanced across trials.

2.2. Results

2.2.1. No-masking: Trial-type analysis

We began our analysis as simply as possible, sorting the data according to trial-type (i.e., based only on the presence or absence of masking dots) independent of the visibility of the masked face. Our first objective was to evaluate whether perceptual averaging occurred on trials in which the masking dots did not linger after the offset of the faces—the no-masking condition. We conducted a repeated-measures 2 (cued face: neutral, emotional) × 2 (arrangement: within-hemifield, between-hemifield) × 2 (emotion-in-pair: happy, angry) analysis of variance (ANOVA) on ratings of the cued face from no-masking trials. As expected, we found a main effect of cued face, $F(1, 29) = 6.56$, $p = .02$, $\eta_p^2 = .184$, and a main effect of emotion-in-pair, $F(1, 29) = 62.19$, $p < .01$, $\eta_p^2 = .682$, but not arrangement $F(1, 29) = 1.24$, *n.s.* The interaction between cued face and emotion-in-pair was significant, $F(1, 29) = 27.14$, $p < .01$, $\eta_p^2 = .483$, but the interaction between cued face and arrangement was not significant, $F(1, 29) = .62$, *n.s.*, nor was the interaction between emotion-in-pair and arrangement, $F(1, 29) = 1.69$, *n.s.* Most important, there was a significant three-way interaction between cued face, arrangement, and emotion-in-pair, $F(1, 29) = 9.56$, $p < .01$, $\eta_p^2 = .248$. This interaction reflects the different effects of the hemifield arrangements on ratings of the cued faces as a function of the emotion in the pair.

Fig. 2A illustrates how ratings of neutral faces were less similar to each other when paired with happy and angry faces in the within-hemfield condition than when paired with the same emotional faces in the between-hemifield condition. Conversely, Fig. 2C illustrates how ratings of happy and angry faces were more similar to each other when they were paired with neutral faces in the within-hemifield condition than in the between-hemifield condition. This pattern of results—*perceptual averaging*—is important for two reasons: first, it replicates the results of Sweeny et al. (2009), and second, it provides a baseline measure for the amount of averaging we should expect in the current investigation independent of any effect of OSM or visual awareness.

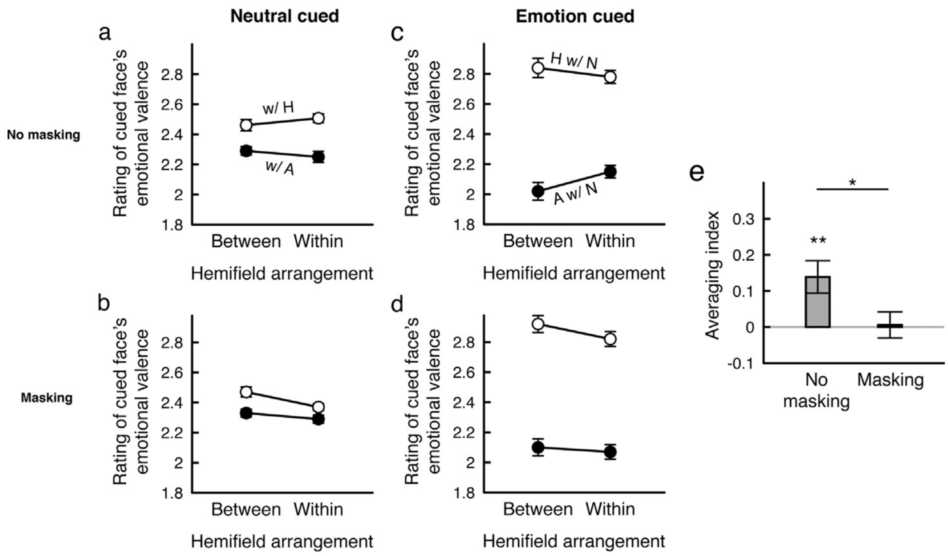

Fig. 2. Effects of perceptual averaging based on trial type (no-masking and masking) and the cued face (neutral and emotional). (A) Emotion ratings (ranging from most negative, 1, to most positive, 4) for a neutral face when an accompanying happy or angry face appeared in the opposite visual hemifield (between-hemifield arrangement) or in the same visual hemifield (within-hemifield arrangement) in the no-masking condition. (B) Emotion ratings for a neutral face accompanied by a happy or angry face in these same hemifield arrangements in the masking condition. (C) Emotion ratings for a happy or angry face when a neutral face appeared in the opposite visual hemifield or within the same hemifield in the no-masking condition. (D) Emotion ratings for a happy or angry face paired with a neutral face in these same arrangements in the masking condition. (E) The averaging index—a metric of the total amount of perceptual integration between the neutral and emotional faces—shown separately for the no-masking and masking trials, collapsed across the emotion of the cued face. The error bars in panels A-D represent ± 1 *SEM* with the baseline individual variability across the four data points removed (i.e., repeated-measures error bars). The error bars in panel E have not been corrected for multiple comparisons across conditions in order to emphasize comparison against a null value of zero.

2.2.2. Perceptual averaging index

The interaction between arrangement (within- vs. between-hemifield) and emotion-in-pair (happy, angry) was the primary focus of this experiment. This two-way interaction was not significant in the repeated measures analysis of variance reported above because it materialized as opposing patterns depending upon which face was cued (see Fig. 2). These opposing patterns reflect the operation of the same underlying mechanism of perceptual averaging, albeit shifting judgments of the cued face's valence in different directions. We were more interested in the strength of perceptual averaging in general, regardless of its direction. We thus sought to translate these interactions into a straightforward metric of perceptual averaging, creating an averaging index to reflect the magnitude of the interaction between arrangement and emotion-in pair. This approach allowed us to simplify comparisons of perceptual averaging as a function of OSM and streamline the presentation of our data. Henceforth, we report results in terms of this averaging index and include results of associated ANOVAs in the Supplemental Materials.

We conceived of the averaging index as a difference of difference scores in which positive values indicated that perceptual averaging had occurred. To calculate the averaging index on trials in which the emotional face was cued (Fig. 2C), we subtracted ratings of happy faces on within-hemifield trials from rating of these same faces on between-hemifield trials (yielding a positive value if averaging had occurred). We then computed a separate difference score for ratings of angry faces, this time subtracting ratings from the between-hemifield condition from ratings from the within-hemifield condition (again yielding a positive value if averaging had occurred). We then added these two difference scores to obtain a total value of averaging on trials in which an emotional face was cued. We computed a separate averaging index for trials in which a neutral face was cued, reversing the subtractions because the overall pattern was expected to go in the opposite direction (see Fig. 2A). We used this approach to obtain averaging indices for each observer.[1]

As expected, the averaging index was positive and significantly different from zero for no-masking trials in which the emotional face was cued, $t(29) = 3.08$, $p < .01$, $d = .56$, though curiously not for trials in which the neutral face was cued $t(29) = 1.37$, *n.s.* The average of these scores provides an overall measure of perceptual averaging on trials with no masking, and this score was significantly different from zero, $t(29) = 3.09$, $p < .01$, $d = .61$ (Fig. 2E). This effect cannot be explained in terms of proximity since spatial distance between faces was greatest in the within-hemifield condition. Nor can it be described as an artifact of rating the uncued face in each pair, since such an error would have been equally likely in both conditions.

2.2.3. Masking: Trial-type analysis

We followed the same procedure described above, this time analyzing data from trials in which the quartet of masking dots was allowed to linger around the locations of one face in each pair. Neither the averaging index for masking trials in which the emotional face was cued, $t(29) = 1.25$, *n.s.* (Fig. 2D), nor the index for trials in which the neutral face was cued, $t(29) = 1.05$, *n.s.* (Fig. 2B), was significantly different from zero. The average of these indices was also not significantly greater than zero, $t(29) = .16$, *n.s.* (Fig. 2E). To determine if the strength of perceptual averaging was weaker in the masking-condition than in the no-masking condition, we conducted a paired-samples *t*-test between the averaging indices from the masking and no-masking conditions. This analysis confirmed that perceptual averaging in the masking condition was indeed significantly weaker than in the no-masking condition, $t(29) = 2.66$, $p = .01$, $d = .49$ (Fig. 2E).

2.2.4. No-masking: Experience-based analyses

In the preceding analyses, trials were categorized as masking or no-masking based solely on the presence of masking dots. Although this kind of approach is favored more often than not in studies of OSM, it does not take into account an observer's visual experience. We therefore repeated the previous series of analyses, this time using each observer's reports of their subjective awareness on a trial-by-trial basis to assign data to the no-masking and masking conditions. This allowed us to evaluate the extent to which perceptual averaging occurred as a function of OSM's effects on visual awareness. We limited the no-masking condition to contain data only from trials in which (a) masking dots did not linger after the offset of the faces and (b) observers reported seeing two faces.

We computed the averaging indices for these no-masking trials and compared them against a null value of zero. Once again, the averaging index was positive and significantly different from zero for trials in which the emotional face was cued, $t(29) = 2.14$, $p = .04$, $d = .39$ (although this test did not survive correction for multiple comparisons using the Benjamini-Hochberg procedure). This time, the averaging index for trials in which the neutral face was cued was also significantly different from zero, $t(29) = 2.71$, $p = .01$, $d = .49$. The composite of these two averaging indices was significantly different from zero as well, $t(29) = 3.10$, $p < .01$, $d = .56$ (see Fig. 3 for a depiction of data from this experience-based analysis).

2.2.5. Masking: Experience-based analyses

Similarly, we re-analyzed the masking condition by including only trials in which (a) masking dots lingered after the faces disappeared, and (b) observers reported seeing only one face. This approach allowed us to look specifically at masking trials in which subjective awareness was disrupted, a luxury that was not possible when sorting data based on trial-type alone. Four observers did not experience the masking effects of OSM, for either some trials types (e.g., within-hemifield masking trials with a neutral face

[1] Henceforth, we report results of statistical analyses on these indices uncorrected for multiple comparisons. We later evaluated our comparisons using the Benjamini-Hochberg procedure (Benjamini, Krieger, & Yekutieli, 2006), pooling all t-tests across Experiments 1 and 2 into a single family for analysis, using a false-discovery rate of 5%. All previously significant comparisons survived this corrective procedure unless explicitly noted.

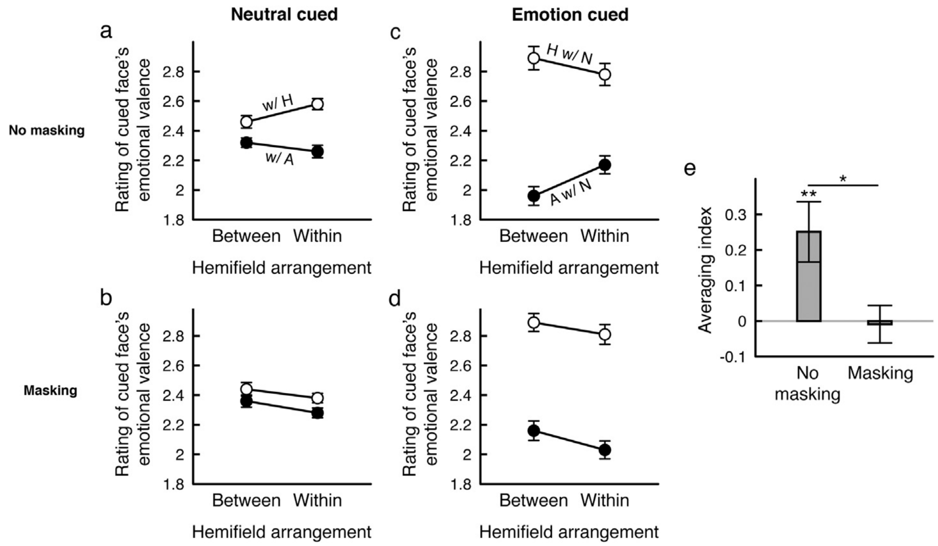

Fig. 3. Effects of perceptual averaging based on visual experience on each trial (no-masking and masking) and the cued face (neutral and emotional) rather than the trial labels alone. (A) Emotion ratings (ranging from most negative, 1, to most positive, 4) for a neutral face when an accompanying happy or angry face appeared in the opposite visual hemifield (between-hemifield arrangement) or in the same visual hemifield (within-hemifield arrangement) in the no-masking condition. (B) Emotion ratings for a neutral face accompanied by a happy or angry face in these same hemifield arrangements in the masking condition. (C) Emotion ratings for a happy or angry face when a neutral face appeared in the opposite visual hemifield or within the same hemifield in the no-masking condition. (D) Emotion ratings for a happy or angry face paired with a neutral face in these same arrangements in the masking condition. (E) The averaging index, shown separately for the no masking and masking trials, collapsed across the emotion of the cued face. The error bars in panels A-D represent ± 1 *SEM* with the baseline individual variability across the four data points removed (i.e., repeated-measures error bars). The error bars in panel E have not been corrected for multiple comparisons across conditions in order to emphasize comparison against a null value of zero.

cued), or all trial types. Since these four observers did not contribute full data to the masking condition when trials were re-categorized by experience, we did not include them in the following analyses.[2] We compared the averaging index from the no-masking condition to that of the masking condition using only data sorted by experience using a paired-samples *t*-test. As with our original analysis by trial type, this analysis based on visual experience again confirmed that perceptual averaging was weaker in the masking condition than in the no-masking condition, $t(25) = 2.18$, $p = .03$, $d = .428$ (Fig. 3E).

2.2.6. Failed masking: Experience-based analyses

We conducted an additional analysis to evaluate the strength of perceptual averaging on trials in which masking dots lingered after the faces disappeared, but observers still indicated that they perceived two faces. On trials in which OSM did not disrupt face detection, the averaging index was positive and significantly different from zero for trials in which the emotional face was cued ($M = 0.329$, $SD = 0.613$), $t(29) = 2.94$, $p = .006$, $d = .537$, but not for trials in which the neutral face was cued ($M = -0.079$, $SD = 0.507$), $t(29) = 0.858$, $p = .397$, $d = .156$. The difference between these indices was also significant, $t(29) = 2.55$, $p = .01$, $d = .465$.

Although these findings may seem surprising, they make sense in the context of previous work which demonstrated that OSM can independently influence detection of an object and discrimination of its features, with discrimination being the easier of the two to eliminate (Gellatly et al., 2006). When observers in our investigation detected the presence of the masked face, they may have missed its expression (assuming it actually displayed happiness or anger). These semi-masked emotional faces could have been perceptually and functionally equivalent to neutral faces, presumably providing no affect for averaging with the cued neutral face. Indeed, perceptual averaging did not occur on these trials when the cued face was neutral, effectively reflecting averaging between two neutral faces. On the other hand, when the masked face was neutral, failure to discriminate it as such would not have been so costly, since it would have been seen merely as a face without affect—more or less equivalent to a neutral face. Accordingly, perceptual averaging still occurred on trials where the cued face was emotional, possibly because the second (masked) face's lack of affect still detracted from that of the cued face.

To summarize, in Experiment 1 we replicated an effect of perceptual averaging when both faces in a pair were visible. More importantly, the emotions from a pair of faces were not averaged when one face received OSM. Additional analyses taking into account the effect of OSM on visual awareness revealed that only the combination of masking dots *and* loss of awareness was enough to significantly reduce averaging relative to the no-masking condition. Averaging always occurred when observers were aware of the

[2] We expected individuals to vary both in their subjective awareness of the cued face and their thresholds for reporting awareness (Sweeny, Suzuki, Grabowecky, & Paller, 2013). It was impossible to predict how much data each observer would contribute to the no-masking and masking conditions after sorting based on visual experience in Experiments 1 and 2. We accepted the likelihood of cutting observers from our analyses as a suboptimal but necessary outcome of our approach.

faces, and averaging did not occur when awareness was eliminated. Thus, perceptual averaging of facial expression is disrupted when one face in a pair is rendered invisible by OSM.

3. Experiment 2

In Experiment 1, we found no evidence that the emotion of a face rendered invisible by OSM could spread to the perception of a nearby face that was clearly visible. This result represents a meaningful advance in understanding visual awareness so long as it cannot be explained by alternative factors that owe nothing to consciousness. For example, if something particular about our stimuli, their timing or placement, or the nature of our response procedure prohibited perceptual averaging in general, then our results in the absence of visual awareness would be uninteresting. Crucially, we *did* replicate a reliable effect of emotion averaging when both faces in a pair were visible (Sweeny et al., 2009, 2011), suggesting that our effects in the masking condition were related to changes in visual awareness. This suggests that the circumstances in which perceptual averaging with faces can be expected to occur are limited. We conducted Experiment 2 to further examine the boundaries of perceptual averaging with faces, in this case when attentional resources were scarce.

Although visual awareness and attention tend to co-occur, they are in fact distinct processes (Webb, Ingelstrom, Shurger, & Graziano, 2016). Interestingly, a growing body of work suggests that the neural representation of emotional facial expressions is not necessarily automatic, and is substantially reduced under situations of high attentional load (e.g., Palermo & Rhodes, 2007; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Tomasik, Ruthruff, Allen, & Lien, 2009; Williams, McGlone, Abbott, & Mattingley, 2005). We thus designed Experiment 2 to feature a dual-task procedure that allowed us to measure the strength of emotion averaging (our primary focus) when attention was simultaneously engaged on the perception of orientation (our secondary focus). We were particularly interested to know whether emotion averaging would persist when faces were visible yet their emotional content was not the exclusive focus of attention.

Experiment 2 also provided an opportunity to evaluate whether an averaging effect with a simpler visual feature—-orientation—could still occur within the context of our experimental design, even in the absence of visual awareness. Demonstrating averaging of head orientation would allow us to make a stronger case that it is the perceptual averaging of emotional expressions, in particular, that does not survive disruptions of awareness from OSM. This outcome is reasonable to expect. When multiple objects are seen at once, and with distributed attention, the visual system tends to pool their orientations into a kind of texture representation even in the absence of crowding (Baldassi & Burr, 2000; Morgan, Ward, & Castet, 1998). Simpler visual properties of objects like orientation (Jacoby, Kamke & Mattingly, 2013) and aspect ratio (Sweeny et al., 2017) even influence perceptual judgments when they are rendered invisible by OSM. We speculated that, when presented on a face, a more basic visual attribute like orientation might influence judgments of other nearby faces even when its visibility is disrupted by OSM.

The averaging effect we demonstrated in Experiment 1 is unique in that it was especially strong when faces were presented within a visual hemifield. This effect occurred on top of a more generic effect of attraction that operated independent of receptive field sizes. For example, even in the between-hemifield condition of Experiment 1, on no-masking trials, neutral faces were rated as more positive when paired with happy faces ($M = 2.41$, $SD = .31$) than when paired with angry faces ($M = 2.26$, $SD = .28$), $t(61) = 4.54$, $p < .01$, $d = .58$. Averaging effects with visual features processed in more basic stages of analysis follow this same pattern, with no dependence on spatial arrangement (Choo & Franconeri, 2010; Jacoby, Kamke & Mattingly, 2013; Sweeny et al., 2017). In fact, orientation averaging seems to engage a kind of "second stage" integration, where information is pooled across the outputs of individual feature detectors, presumably with smaller receptive fields (Baldassi & Burr, 2000). When examining orientation averaging in Experiment 2, we thus expected a general effect of orientation averaging to occur regardless of spatial arrangement. We also note, however, that neurophysiological work has documented the existence of cells in relatively high levels of the ventral visual pathway (e.g., IT) that respond selectively to head tilt (Ashbridge, Perrett, Oram, & Jellema, 2000; Freiwald & Tsao, 2010; Tanaka, Saito, Fukada, & Moriya, 1991). We were thus open to the possibility that an additional effect of hemifield averaging might occur for the perception of face orientation.

3.1. Material and methods

3.1.1. Observers

We expected Experiment 2 to be more difficult that Experiment 1 by virtue of its dual-task design, and to therefore produce noisier data. We thus approximately doubled the number of observers recruited for Experiment 2. Sixty-three undergraduate students from the University of Denver participated for course credit, each completing 560 trials. None of the subjects in Experiment 2 participated in Experiment 1, and each was naïve regarding the purpose of the investigation. The number of trials per observer was reduced relative to Experiment 1 so that each observer could complete all trials within an hour and a half. Observers granted informed consent and had normal or corrected-to-normal vision.

3.1.2. Stimuli

Stimuli were identical to those used in Experiment 1 with the exception that, on every trial, each face in a pair was independently tilted to the left or right (i.e., in a counter-clockwise or clockwise direction) (Fig. 4A). Relative to the 12o'clock upright position, we refer to counter-clockwise orientations as negative and clockwise orientations as positive. On each trial, the orientations of each face in a pair were independently selected from a uniform distribution with a randomly selected mean between -15° and $+15^{\circ}$ and a width of 40°. For example, if the mean of the sampling distribution were -10° on a given trial, the orientations of each face could

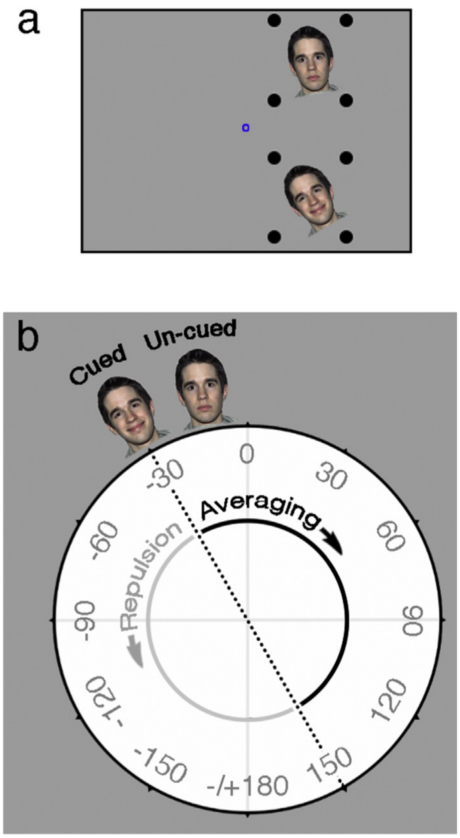

Fig. 4. (A) A typical face arrangement from an orientation trial in Experiment 2. (B) The orientation of each face was randomly determined on each trial. Face orientations never exceeded −35° or +35°, and the difference between the orientations within a pair was free to vary by as much as 40°. See the Stimuli section for more details. On each trial, we used an averaging index to quantify the extent to which the observer's judgment of the cued face's orientation was pulled toward direction of the un-cued face (averaging), or pushed in the opposite direction (repulsion). Scores in the direction of the un-cued face up to 179° away from the cued face were scored positively, as averaging.

have been between −30° and +10°. This sampling procedure constrained the range of possible orientations to occur between −35° and +35° across the experiment. We selected this limited range based on an anecdotal observation that natural head tilts tend to occur approximately within these values. This sampling procedure also ensured that we would obtain a variety of differences in head tilt across trials (e.g., the difference between the orientations of heads in a pair could have been as small as 0° or as large as 40°). This allowed us to evaluate the extent to which a potential effect of orientation averaging depended on the difference between the orientations in each face pair.

3.1.3. Procedure

The procedure for Experiment 2 was identical to that of Experiment 1 with the following exceptions. Trials were randomly determined to be either an *expression trial* or an *orientation trial*, depending on which qualities of the target face were to be rated. On expression trials, the response procedure was identical to that used in Experiment 1. Immediately after orientation trials, however, observers viewed a different response screen prompting them to "Indicate the degree of orientation on the cued image" and asking "How many faces did you see clearly?". A response face with a neutral expression appeared at the center of this display with an initial orientation randomly sampled from a uniform distribution between −35° and 35°. The identity of the response face was randomly selected. Observers used one of four keys on the keyboard to adjust the orientation of the response face to match the orientation of the cued face. These four keys overlapped with those used to respond to expression trials. One key rotated the face counter-clockwise by −10°, another by −1°, another by +1°, and another by 10° clockwise. Observers simply adjusted the face until they were satisfied and then pressed the left or right arrow to indicate the number of faces that were visible during the trial, at which point the trial ended. The final orientation of the response face was recorded in terms of degrees (+ or −) relative to perfectly upright. Note that observers could not predict which aspect of the cued face (emotion or orientation) they would be asked to evaluate. This should have encouraged them to attend to both the emotion and orientation of the faces.

Observers had unlimited time to respond. All 64 observers completed 560 trials, which were counterbalanced across the within- and between-hemifield spatial arrangements. The emotional expression (happy or angry) paired with the neutral face, the location of the emotional face (top, bottom, left, or right), and the cued face (neutral or emotional) were also counterbalanced across trials.

3.2. Results

Data from two observers were cut as a result of failure to follow the instructions. Below, we first present results from emotion-related analyses (i.e., ignoring data from orientation trials). Then we present results from orientation-related analyses.[3]

3.2.1. No-masking: Trial-type emotion analysis

We again began by sorting the data according to trial-type, based on the presence or absence of masking dots independent of the experience of the observer. We computed the perceptual averaging index for Experiment 2 in the same manner as Experiment 1. This time, however, the averaging index was not significantly different from zero for no-masking trials in which the emotional face was cued, ($M = .006$, $SD = .4$), $t(61) = .12$, *n.s.*, nor for trials in which the neutral face was cued, ($M = .007$, $SD = .35$), $t(61) = .15$, *n.s.* The average of these scores was not significantly different from zero, $t(61) = .89$, *n.s.* In other words, perceptual averaging of emotional expressions did not occur in Experiment 2.

3.2.2. Masking: Trial-type emotion analysis

We followed the same procedure described above, this time analyzing data from trials in which a quartet of masking dots was allowed to linger around the location of one face in each pair. Neither the averaging index for masking trials in which the emotional face was cued, ($M = .01$, $SD = .64$), $t(61) = .14$, *n.s.*, nor the index for trials in which the neutral face was cued, ($M = .009$, $SD = .39$), $t(61) = .18$, *n.s.*, was significantly different from zero. The average of these indices was also not significantly greater than zero, $t(61) = .63$, *n.s.* In other words, we found no evidence of perceptual averaging when one face in a pair was surrounded by masking dots.

3.2.3. No-masking: Experience-based emotion analyses

We again repeated the preceding series of analyses, this time using each observer's reports of their subjective awareness on a trial-by-trial basis to assign data to the no-masking and masking conditions, in the same manner as Experiment 1. We began with no-masking trials. Notably, two observers never reported seeing both faces. These two observers were cut from the following analyses. We computed the averaging indices for no-masking trials, according to subjective experience, and compared them against a null value of zero. The averaging index was not different from zero for trials in which the emotional face was cued, ($M = -.06$, $SD = .45$), $t(59) = 1.01$, *n.s.*, nor for trials in which the neutral face was cued, ($M = -.006$, $SD = .42$), $t(59) = .1$, *n.s.* The composite of these two averaging indices was not significantly different from zero, $t(59) < .01$, *n.s.*.

3.2.4. Masking: Experience-based emotion analyses

Similarly, we re-analyzed the masking condition by including only trials in which (a) masking dots lingered after the faces disappeared, and (b) observers reported seeing only one face. Although the overwhelming majority of observers experienced disruptions of visual awareness from OSM, only thirty-one observers reported disruptions with each combination of condition and type of stimulus (e.g., within-hemifield masking trials with a neutral face cued, between-hemifield masking trials with an angry face cued, etc.). We only intended to analyze data from observers who contributed full data sets, so we cut the remaining 31 observers who did not meet this requirement from the following analyses. Neither the averaging index for trials in which the emotional face was cued, ($M = -.08$, $SD = 1.73$), $t(30) = .47$, *n.s.*, nor the averaging index for trials in which the neutral face was cued, ($M = -.03$, $SD = 1.6$), $t(30) = .28$, *n.s.*, was significantly different from zero.

The demanding nature of Experiment 2 may have disrupted the deployment of attentional resources for encoding the emotion on the faces. If this were true, then there would have been less emotional signal for the visual system to average. Indeed, the difference between ratings of happy and angry faces in Experiment 2 (collapsed across spatial arrangements in the no-masking experience based analysis; M = .4, SD = .46) was smaller than the difference in Experiment 1 (M = .77, SD = .73). In this sense, our results most likely reflect a disruption in the initial encoding of emotion and a failure of perceptual averaging to act on these weaker emotional signals (we rule out alternative explanations later, in the Discussion section).

3.2.5. Orientation averaging index

Next, we sought to determine whether perceptual averaging occurred on orientation trials. In particular, we aimed to determine whether perception of the cued face's orientation was pulled toward the orientation of the un-cued face. For example, suppose a cued face was rotated -30° and an un-cued face was rotated -10°. A response of $-25°$ would indicate an effect of averaging in which perception of the cued face's orientation was "pulled" 5° in the direction of the un-cued face's orientation. A response in the opposite direction would indicate an effect of perceptual repulsion (Fig. 4B).

Similar to emotion trials, we created an orientation averaging index that indicated both the direction (attraction or repulsion) and magnitude of orientation distortion on a trial-by-trial basis. In this index, positive values reflected orientation averaging and negative values reflected repulsion. Because observers indicated the orientation of the cued face by rotating a response face 360° around its center, averaging index values could have been up to $\pm$ 180°. Values of exactly $\pm$ 180° relative to the cued face were excluded since it

[3] Due to a programming error, for emotion-trials only, observers were sometimes able to advance past the response screen after indicating the number of faces visible without reporting the emotion of the cued face. These trials, which amounted to 1.8% of emotion trials per observer, on average, were eliminated from all analyses.

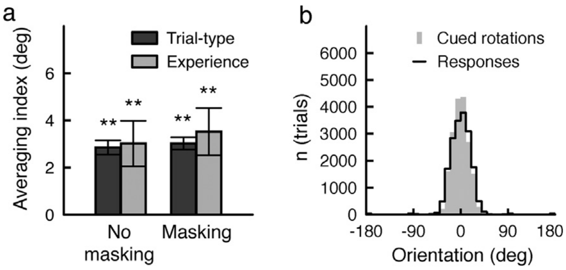

Fig. 5. (A) Effects of orientation averaging with and without masking based on the presence of masking dots (trial-type) and the observer's visual awareness of the cued face (experience). (B) The distribution of orientations of the cued face (filled gray histogram) and the distribution of orientations selected by observers (open black histogram). Histograms reflect pooled data from all conditions and all trials from all observers. Error bars in panel A represent ± 1 *SEM*, uncorrected for multiple comparisons across conditions in order to emphasize comparison against a null value of zero.

was unclear whether this constituted averaging or repulsion (this occurred on fewer than .5% of all trials).

3.2.6. Trial-type orientation analyses

We again began our analyses as simply as possible by sorting the data according to trial-type, based on the presence or absence of masking dots independent of the experience of the observer. We calculated each observer's mean orientation averaging index across trials in which masking dots were not present. A one-sample *t*-test confirmed that, across observers, the orientation averaging index was positive and different from zero, ($M = 2.85°$, $SD = 2.35°$), $t(61) = 9.57$, $p < .01$, $d = 1.22$ (Fig. 5A). Additionally, the orientation averaging index for no-masking within-hemifield trials ($M = 2.7°$, $SD = 2.71°$) was smaller than it was for between-hemifield trials ($M = 3.34°$, $SD = 2.11°$), although this difference was not significant $t(61) = 1.91$, $p = .061$. The orientation index was also positive and different from zero for trials in which masking dots were present, ($M = 3.02°$, $SD = 2.04°$), $t(61) = 11.67$, $p < .01$, $d = 1.48$ (Fig. 5A). The orientation averaging index for within-hemifield trials ($M = 2.7°$, $SD = 2.71°$) was again smaller than it was for between-hemifield trials ($M = 3.33°$, $SD = 2.12°$), but again this difference was just a trend, $t(61) = 1.92$, $p = .06$. Overall, the difference between the orientation averaging index on no-masking and masking trials was not significant, ($M = .17°$, $SD = 1.88°$), $t(61) = .73$, *n.s.*

3.2.7. Experience-based orientation analyses

Next, we calculated the orientation averaging index for trials in which (a) masking dots did not linger after the offset of the faces and (b) observers reported seeing two faces—the *no-masking condition*. The index was positive and different from zero, $t(61) = 3.14$, $p < .01$, $d = .40$ (Fig. 5A). This time, however, the orientation averaging indices for within-hemifield trials ($M = 2.68°$, $SD = 2.44°$) and for between-hemifield trials ($M = 3.62°$, $SD = 6.48°$) were not significantly different $t(59) = .3$, *n.s.* Therefore, in all analyses henceforth, we collapse data across between and within-hemifield trials.

The mean orientation averaging index for trials in which trailing dots lingered after the offset of the faces and observers reported seeing just one face—the *masking condition*—was also positive and different from zero, $t(52) = 3.52$, $p < .01$, $d = .48$ (Fig. 5A). Note that nine observers reported seeing both faces on all orientation trials and were excluded from the preceding analysis. Looking only at observers who produced data for both conditions, the difference between the averaging index from the no-masking and masking conditions according to experience was not significant, ($M = .28°$, $SD = 10.37°$), $t(52) = .20$, *n.s.* Thus, orientation averaging between faces in a pair was always strong, regardless of whether both faces were visible.

3.2.8. Averaging as a function of orientation difference

Next, we conducted an exploratory analysis to determine whether perceptual averaging varied according to the difference between the orientations of the two faces in each pair. Previous work showed that attractive interactions between simultaneously presented shapes increase in strength as the shapes in a pair become more distinct (Sweeny et al., 2017). We expected a similar type of effect here in which orientation averaging would be stronger between pairs of faces with relatively disparate orientations (e.g., $-20°$ and $+20°$), compared to pairs of faces with relatively similar orientations (e.g., $-5°$ and $+5°$). To address this question, we began by sorting trials based on the presence or absence of masking dots as well as the phenomenology of observers, into no-masking trials and masking trials, separately for each observer. Then, for each observer, we plotted each trial's averaging index against the orientation difference between the faces in the pair. We then obtained a linear fit across the resulting scatterplot for each observer, extracting the slope of the linear fit as an index of the extent to which attraction depended on the difference in orientation between the faces. We expected slopes to be positive overall, and significantly greater than zero in one-sample *t*-tests.

For no-masking trials based on experience, the mean slope was positive and greater than zero, indicating that averaging increased in strength as faces became more disparate in terms of orientation, $t(60) = 6.39$, $p < .01$, $d = .82$ (two observers reported seeing both faces on no-masking trials only once, which prevented those observers from contributing slopes to the analysis).

We then turned to masking trials in which observers reported seeing only one face. Thirteen observers failed to meet this requirement, which prevented them from contributing slopes to the analysis below. Six additional observers experienced disruptions of visual awareness on fewer than ten total trials, compared to the group average of about thirty-seven. We were concerned about the reliability of linear fits applied to the scarce data from these six observers so we cut them from our initial analysis (although we conducted additional bootstrapping tests that include their data).[4] Across the remaining forty-four observers, the mean slope was positive and different from zero $t(43) = 3.89$, $p < .01$, $d = .59$. This suggests that the more disparate in terms of orientation the two faces were, the more orientation averaging occurred, even if only one face was consciously experienced. We note that this pattern of results differs from those which typically occur during crowding (e.g., averaging turning into repulsion, Hariharan, Levi, & Klein, 2005), but because the effect we have observed here is not the result of crowding, per se, it should not necessarily be expected to produce the same results.

3.2.9. Ruling out alternative explanations

Observers performed the orientation task quite well; the correlation between cued orientation and response orientation was strong (*Mean* $R = 0.66$, $SD = 0.19$), and the distribution of raw orientation responses across all observers was narrow and similar to the distribution of the cued face orientations (Fig. 5B). This suggests that if observers did guess, they did so infrequently and within a highly restricted range. Nevertheless, it is reasonable to expect that observers responded randomly on some trials, or used alternative response strategies, and to consider whether each could have produced a pattern of results superficially similar to our findings of orientation averaging. Below we discuss some alternative explanations and the extent to which they constrain our interpretations.

First, rating the un-cued face (either intentionally or by accident) would have produced an averaging index that was both positive and increasing in magnitude with greater differences between the two faces' orientations (i.e., positive slopes as described above). This possibility is very unlikely, though, because we still observed orientation averaging even when the un-cued face was not visible.

Second, pseudo-random guessing could have produced a pattern of results very similar to what we observed, albeit exaggerated in magnitude. We verified this by constructing four simulated data sets, each containing 100,000 trials with sampling of head-pair orientations identical to that from our actual experiment. In these four simulated sets, we randomly sampled responses from (a) the entire 360° range of orientations, (b) the upright range, $-90°$ to 90°, (c) the actual range of head orientations, $-35°$ to 35°, and (d) a normal distribution with a mean (0.25°) and *SD* (14.8°) identical to that of the actual distribution of orientation responses from Experiment 2 (see Supplemental Section for additional details about these simulations). These simulations confirmed that random responding within the upright range (simulations b-d) would have led to a positive averaging index overall, and crucially, would *always* have led to positive slopes. Consistent with this idea that positive slopes were related to guessing, we also found a negative relationship between slope and task performance (indexed by the correlation between cued rotation and response, described above) in our data, ($R = -0.65$). That is, observers who produced the least accurate data also produced the steepest slopes.

Although the results of these simulations limit the extent to which we can interpret the positive slopes we reported in the preceding analyses, they also suggest something far more useful—any observers who produced *negative* slopes could not have been guessing, at least not enough to distort their overall patterns of data. We thus conducted a conservative analysis of orientation averaging, looking only at these fourteen observers who produced negative slopes (and not coincidentally, strong correlations between cued orientation and response orientation, *Mean* $R = 0.78$, $SD = 0.16$). Among these high-quality observers, we still found evidence of orientation averaging (in the no-masking condition when both faces were visible, $M = 1.2$, $SD = 1.1$, $t[13] = 4.07$, $p < .01$, $d = 1.09$; and a trend in the masking condition when only one face was visible, $M = 2.3$, $SD = 4.05$, $t[11] = 1.97$, $p = .07$, $d = 0.57$).

Finally, the disappearance of the averaging effect with regard to emotional expression could have simple been because observers put little effort into the expression task, and instead directed all their attention to the orientation task, presumably because it was easier. To evaluate this less interesting possibility, we examined the extent to which strong performance on the orientation task depended on poor performance on the emotion task. We again quantified each observer's orientation performance as the correlation between the actual orientation of the cued face and their orientation response on each no-masking trial. To index performance on the emotion task, we measured the difference between the average rating of happy faces and angry faces on emotion-cued trials from the no-masking condition. In this case, higher difference scores would indicate better recognition of the emotion on the cued faces. If stronger orientation performance came at a cost of emotion recognition, these two indices should have been negatively correlated. Instead, performance on the orientation task was positively correlated with emotion recognition ($R = 0.21$, $p = 0.1$).

In summary, in Experiment 2 we demonstrated that perceptual averaging of facial expression does not occur when attention is occupied by a demanding secondary task, even if both faces in a pair are visible. In contrast, orientation averaging appears to occur with limited attentional resources and even when one face in a pair is not visible. However, we hesitate to make conclusions about the *exact* extent to which this orientation averaging reflected a pure change in perception or possibly some combination of a perception distortion and random responding.

4. Discussion

We evaluated the extent to which face-specific emotion averaging—a process in which two faces appear more like each other

[4] Cutting these six observers should not have systematically skewed our evaluation of slopes at the group level. For trials in which masking dots were present and one face was visible, across all observers, the slopes of the linear fits were not related to the number of trials from each observer, $R(49) = .07$, $p = .63$. Thus, if anything, removing these observers only eliminated noise from our analysis. We also conducted a bootstrapping analysis to evaluate slopes in a way that included data from these six observers. Specifically, we combined all masking trials ($N = 2378$) in which our observers reported seeing just one face into a single distribution. Then, we sampled from this distribution (with replacement) to construct a bootstrapped distribution of averaging scores and orientation differences. We did this 10,000 times, applying a linear fit to each bootstrapped distribution and then obtaining the resulting slopes. The slopes from our masking trials were greater than the slopes of a null-distribution obtained using Monte Carlo methods (*bootstrapped p-value* $< .01$).

when they are seen at the same time—can occur under conditions of limited visual awareness and attention. We replicated previous work showing that this process occurs when both faces are visible and are the exclusive focus of attention (Sweeny et al., 2009), but crucially, this process is disrupted when one face is rendered invisible by OSM, or when attention is engaged by a secondary task. Broadly speaking, these results fit with the notion that information integration is limited without consciousness (Mudrik, Faivre, & Koch, 2014), especially when that information is visually complex and is processed relatively late by the visual system.

These results are important for several reasons. First, although previous work using fMRI (Carlson et al. 2007) might hint at the perceptual outcome we observed, it was also possible that lingering activity in face-processing areas—activity not captured by physiological measures—could persist during OSM and could influence perception. This does not seem to be the case. Notably, our results do not necessarily mean that face representation is completely eliminated under OSM. In fact, there is some evidence that macaque STS can remain active, even when viewing a masked face (e.g., Rolls & Tovee, 1994; Rolls, Tovee, Purcell, Stewart & Azzopardi, 1994). We more cautiously suggest that if there was any lingering representation of facial expression in our investigation, it was incapable of influencing the perception of neighboring visible faces. Secondly, our work dovetails with other investigations to suggest that even though averaging of relatively simple features can occur in the absence of awareness during OSM (Choo & Franconeri, 2010; Jacoby, Kamke & Mattingly, 2013; Sweeny et al., 2017), the same cannot be expected for complex information like facial expression. This is in line with the idea that visual processing and perception become more strongly related to one another at later, more complex stages of visual analysis. We also caution that this relationship does not necessarily imply causality—disrupted averaging and disrupted awareness may simply be correlated under OSM.

Experiment 2 featured a dual-task design in which observers had to attend to both the expressions and orientations of faces. Expression averaging was eliminated in this new context, even when both faces were clearly visible. This outcome did not occur as a result of observers simply giving up on the expression task, nor did it occur because rotating the heads made expression recognition more difficult (see Supplemental section). It was also very unlikely to reflect a simple failure to replicate—Experiment 1 produced a pattern of results nearly identical to those from a previous investigation with a different set of faces (Sweeny, Grabowecky, Paller, & Suzuki, 2009). Instead, it seems more reasonable that the increased attentional demand of Experiment 2 disrupted the representation of emotional expression and associated averaging. This account is in line with work which suggests that the processing of emotional faces requires attention (Pessoa et al., 2002; Williams et al., 2005; Tomasik et al., 2009).

Nonetheless, even when observers were only subjectively aware of one face and emotion averaging did not proceed, some information about the orientation of each face was processed, averaged, and then incorporated into perceptual judgments. Although, as discussed above, averaging of simple features under OSM has been demonstrated before, most previous work has not evaluated effects of awareness on averaging of relatively simple and complex features within the *same* object. In contrast, we have demonstrated that even when extracted from the same object, relatively simple visual information is averaged in the absence of awareness, while more complex information is not. We acknowledge that the expressions we used were not perfectly matched to the orientations in terms of discriminability, but they were nonetheless strong displays of emotion sufficient for averaging in Experiment 1. We suspect that similar results would hold for more precisely matched combinations of simple and complex visual features. At any rate, this result is a novel demonstration that OSM can selectively disrupt interactions between some features of objects while leaving interactions between other features intact.

If OSM does isolate an early stream of feedforward processing (a feature of both substitution and updating accounts of OSM; Di Lollo et al., 2000; Jannati et al., 2013; Kotsoni et al., 2007), then our results suggest that this initial wave of visual analysis is sufficient for orientation averaging, but insufficient for face-specific emotion averaging. This may indicate decaying representation as information travels through successive stages of visual analysis, possibly via the accumulation of neural noise (Faisal et al., 2008). If, on the other hand, OSM operates by interfering with feedforward activation (Macknik & Martinez-Conde, 2007; Põder, 2012), our results could instead reflect more opportunities for inhibition at each stage of visual analysis.

The perceptual averaging of facial expressions we showed in Experiment 1 was not simply the result of crowding. Although crowding does bear some superficial resemblance to our effects (e.g., it is weaker across the vertical meridian than across the horizontal meridian [Liu, Jiang, Sun, & He, 2009]; additionally, see Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), crowding would have been very unlikely given the large inter-stimulus distances in our investigation (e.g., Levi, 2014; Petrov & Meleshkevich, 2011). Nor can our results be accounted for by response biases or strategies, like accidentally rating the un-cued face. Rating the un-cued face would have been equally probable when face pairs were arranged within- and between-hemifields. Rather, we observed stronger perceptual averaging in the within-hemifield condition. This pattern of data is consistent, instead, with neural averaging that occurs when multiple objects fall with the large, hemifield-sized receptive fields of ventral visual neurons that encode information about complex objects (e.g., Boussaoud et al., 1991; Chelazzi et al., 1998; Desimone & Gross, 1979; Niemeier et al., 2004; Op De Beeck & Vogels, 2000). The orientation averaging we showed in Experiment 2 also bears the hallmarks of a true perceptual effect. Although some random responding probably occurred during judgments of face orientation, we showed that this cannot completely account for our results. Nor was orientation averaging the result of rating the un-cued face—it occurred even when the second face from the pair was not even visible. Orientation assimilation can occur in uncrowded displays (Baldassi & Burr, 2000; Morgan, Castet, & Ward, 1998), and our results are most consistent with a simple pooling mechanism that may operate when attention is distributed equally across multiple items in a set.

It is unclear at which stage of visual analysis head orientation was processed in our experimental design. It is plausible that these computations occurred at least partially outside V1 within the human homologue of area of IT, which is known to contain cells sensitive to head tilt (Tanaka et al., 1991; Perrett, Oram, & Ashbridge, 1998; Ashbridge et al., 2000; Freiwald, Tsao, & Livingstone, 2009; Freiwald & Tsao, 2010). These head-orientation selective cells appear to be concentrated in anterior-IT, however more generic shape-tuned cells located in posterior-IT can encode an object's orientation as well (Tanaka et al., 1991). So, our finding of orientation averaging is not necessarily an indication that face-selective processing occurred. Although it is unclear exactly where and when head orientation is processed, we reason that head orientation is a feature likely to at least partially recruit orientation processing areas prior to IT. And, regardless of just how "simple" head tilt is, it is a feature that was averaged during disruptions of visual awareness from OSM, while emotion information was not.

Averaging is a general principle of vision. It occurs both at the level of individual cells (e.g., Kastner et al., 2001; Miller et al., 1993; Rolls & Tovee, 1995; Sato, 1989; Zoccolan, Cox, & DiCarlo, 2005) and across populations of neurons (e.g., Suzuki, Clifford & Rhodes, 2005). Averaging, and the pooling that underlies it, is also central to several perceptual phenomena, including ensemble coding (e.g., Alvarez, 2011; Whitney et al., 2014), crowding (e.g., Whitney & Levi, 2011, Fischer & Whitney, 2011), and the effects demonstrated here. It is unclear to what extent these phenomena may rely on shared mechanisms, but interestingly, each produces a common outcome of perceptual homogeneity. This investigation converges with several others to illustrate that, as important as it may be for this illusion of uniformity to be maintained by the visual system, it does occur with constraints. Whereas visual awareness and attention seem to have little effect on perceptual averaging with relatively simple visual features like orientation, they appear to be important or even necessary for perceptual averaging to proceed with more complex visual information, like facial expression.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.concog.2018.03.005.

References

Adams, W. J., Gray, K. L., Garner, M., & Graf, E. W. (2010). High-level face adaptation without awareness. *Psychological Science, 21*(2), 205–210.
Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122–131.
Argyropoulos, I., Gellatly, A., Pilling, M., & Carter, W. (2013). Set size and mask duration do not interact in object-substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 646–661.
Ashbridge, E., Perrett, D. I., Oram, M. W., & Jellema, T. (2000). Effect of image orientation and size on object recognition: Responses of single units in the macaque monkey temporal cortex. *Cognitive Neuropsychology, 17*(1–3), 13–34.
Axelrod, V., Bar, M., & Rees, G. (2015). Exploring the unconscious using faces. *Trends in Cognitive Sciences, 19*(1), 35–45.
Baldassi, S., & Burr, D. C. (2000). Feature-based integration of orientation signals in visual search. *Vision Research, 40*(10), 1293–1300.
Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika, 93*(3), 491–507.
Boehler, C. N., Schoenfeld, M. A., Heinze, H. J., & Hopf, J. M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences, 105*(25), 8742–8747.
Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *Journal of Comparative Neurology, 306*(4), 554–575.
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.
Bridgeman, B. (2006). Contributions of lateral inhibition to object substitution masking and attention. *Vision Research, 46*(24), 4075–4082.
Carlson, T. A., Rauschenberger, R., & Verstraten, F. A. (2007). No representation without awareness in the lateral occipital cortex. *Psychological Science, 18*(4), 298–302.
Chakravarthi, R., & Cavanagh, P. (2009). Recovery of a crowded object by masking the flankers: Determining the locus of feature integration. *Journal of Vision, 9*(10), 1–9 4.
Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology, 80*(6), 2918–2940.
Choo, H., & Franconeri, S. L. (2010). Objects with reduced visibility still contribute to size averaging. *Attention, Perception, & Psychophysics, 72*(1), 86–99.
Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.). *The handbook of evolutionary psychology* (pp. 584–627). Hoboken, NJ: John Wiley & Sons.
De Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience, 14*(2), 199–209.
Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research, 178*(2), 363–380.
Di Lollo, V. (2014). Reentrant processing mediates object substitution masking: Comment on Põder (2013). *Frontiers in Psychology, 5*(819), 1–5.
Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*(4), 481–507.
Dumoulin, S. O., & Hess, R. F. (2007). Cortical specialization for concentric shape processing. *Vision Research, 47*(12), 1608–1613.
Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble perception of dynamic emotional groups. *Psychological Science, 28*(2), 193–203.
Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research, 44*(12), 1321–1331.
Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8*(2), 135–139.
Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews. Neuroscience, 9*(4), 292–303.
Faivre, N., Berthet, V., & Kouider, S. (2012). Nonconscious influences from emotional faces: A comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Psychology, 3*(129), 1–13.
Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2014). Size (mostly) doesn't matter: The role of set size in object substitution masking. *Attention, Perception, & Psychophysics, 76*(6), 1620–1629.
Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2015). Object substitution masking for an attended and foveated target. *Journal of Experimental Psychology: Human Perception and Performance, 41*(1), 6–10.
Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology, 106*(3), 1389–1398.
Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science, 330*(6005), 845–851.
Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience, 12*(9), 1187–1196.
Gellatly, A., Pilling, M., Cole, G., & Skarratt, P. (2006). What is being masked in object substitution masking? *Journal of Experimental Psychology-Human Perception and Performance, 32*(6), 1422–1434.
Goodhew, S. C. (2017). What have we learned from two decades of object-substitution masking? Time to update: Object individuation prevails over substitution. *Journal of Experimental Psychology: Human Perception and Performance, 43*(6), 1249–1262.
Goodhew, S. C., Edwards, M., Boal, H. L., & Bell, J. (2015). Two objects or one? Similarity rather than complexity determines objecthood when resolving dynamic input. *Journal of Experimental Psychology: Human Perception and Performance, 41*(1), 102–110.
Goodhew, S. C., & Edwards, M. (2016). Object individuation is invariant to attentional diffusion: Changes in the size of the attended region do not interact with object-substitution masking. *Cognition, 157*, 358–364.

Goodhew, S. C., Dux, P. E., Lipp, O. V., & Visser, T. A. (2012). Understanding recovery from object substitution masking. *Cognition, 122*(3), 405–415.
Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*, R751–R753.
Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 718–734.
Haberman, J., & Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychonomic Bulletin & Review, 18*(5), 855–859.
Hariharan, S., Levi, D. M., & Klein, S. A. (2005). "Crowding" in normal and amblyopic vision assessed with Gaussian and Gabor C's. *Vision Research, 45*, 617–633.
Harris, J. A., Ku, S., & Woldorff, M. G. (2013). Neural processing stages during object-substitution masking and their relationship to perceptual awareness. *Neuropsychologia, 51*(10), 1907–1917.
Harris, J. A., Wu, C. T., & Woldorff, M. G. (2011). Sandwich masking eliminates both visual awareness of faces and face-specific brain activity through a feedforward mechanism. *Journal of Vision, 11*(7), 1–12 3.
Harrison, G. W., Rajsic, J., & Wilson, D. E. (2016). Object-substitution masking degrades the quality of conscious object representations. *Psychonomic Bulletin & Review, 23*(1), 180–186.
Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32*(3), 203–218.
Haynes, J. D., Driver, J., & Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron, 46*(5), 811–821.
Hung, Y., Smith, M. L., Bayle, D. J., Mills, T., Cheyne, D., & Taylor, M. J. (2010). Unattended emotional faces elicit early lateralized amygdala–frontal and fusiform activations. *Neuroimage, 50*(2), 727–733.
Jacoby, O., Kamke, M. R., & Mattingley, J. B. (2013). Is the whole really more than the sum of its parts? Estimates of average size and orientation are susceptible to object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 233–244.
Jannati, A., Spalek, T. M., & Di Lollo, V. (2013). A novel paradigm reveals the role of reentrant visual processes in object substitution masking. *Attention, Perception, & Psychophysics, 75*(6), 1118–1127.
Juan, C. H., & Walsh, V. (2003). Feedback to V1: A reverse hierarchy in vision. *Experimental Brain Research, 150*(2), 259–263.
Kahan, T. A., & Enns, J. T. (2010). Object trimming: When masking dots alter rather than replace target representations. *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 88–102.
Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology, 86*(3), 1398–1411.
Keysers, C., Xiao, D. K., Földiák, P., & Perrett, D. I. (2001). The speed of sight. *Journal of Cognitive Neuroscience, 13*(1), 90–101.
Kotsoni, E., Csibra, G., Mareschal, D., & Johnson, M. H. (2007). Electrophysiological correlates of common-onset visual masking. *Neuropsychologia, 45*(10), 2285–2293.
Lamme, V. A., Super, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology, 8*(4), 529–535.
Levi, D. (2014). Visual crowding. In J. S. Werner, & L. M. Chalupa (Eds.). *The New Visual Neurosciences* (pp. 681–694). MIT Press.
Liu, T., Jiang, Y., Sun, X., & He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Current Biology, 19*(2), 127–132.
Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: Using apparent motion to isolate the object-level component of object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 106–120.
Macknik, S. L., & Martinez-Conde, S. (2007). The role of feedback in visual masking and visual processing. *Advances in Cognitive Psychology, 3*(1–2), 125–152.
Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Research, 616*(1), 25–29.
Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1171–1180.
Morgan, M. J., Ward, R. M., & Castet, E. (1998). Visual search for a tilted target: Tests of spatial uncertainty models. *The Quarterly Journal of Experimental Psychology: Section A, 51*(2), 347–370.
Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences, 18*(9), 488–496.
Niemeier, M., Goltz, H. C., Kuchinad, A., Tweed, D. B., & Vilis, T. (2004). A contralateral preference in the lateral occipital area: Sensory and attentional mechanisms. *Cerebral Cortex, 15*(3), 325–331.
Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *Journal of Comparative Neurology, 426*(4), 505–518.
Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*(1), 75–92.
Pallett, P. M., & Meng, M. (2013). Contrast negation differentiates visual pathways underlying dynamic and invariant facial processing. *Journal of Vision, 13*(14), 13.
Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*(7), 739–744.
Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292*(5516), 510–512.
Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition, 67*(1), 111–145.
Petrov, Y., & Meleshkevich, O. (2011). Asymmetries and idiosyncratic hot spots in crowding. *Vision Research, 51*(10), 1117–1123.
Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, 99*(17), 11458–11463.
Pilling, M., Gellatly, A., Argyropoulos, Y., & Skarratt, P. (2014). Exogenous spatial precuing reliably modulates object processing but not object substitution masking. *Attention, Perception, & Psychophysics, 76*(6), 1560–1576.
Põder, E. (2012). Attentional gating models of object substitution masking. *Journal of Experimental Psychology: General, 142*(4), 1130–1141. http://dx.doi.org/10.1037/a0030575.
Prime, D. J., Pluchino, P., Eimer, M., Dell'Acqua, R., & Jolicœur, P. (2011). Object-substitution masking modulates spatial attention deployment and the encoding of information in visual short-term memory: Insights from occipito-parietal ERP components. *Psychophysiology, 48*(5), 687–696.
Reiss, J. E., & Hoffman, J. E. (2007). Disruption of early face recognition processes by object substitution masking. *Visual Cognition, 15*(7), 789–798.
Rolls, E. T., & Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society of London-B-Biological Sciences, 257*(1348), 9–16.
Rolls, E. T., & Tovee, M. J. (1995). The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Experimental Brain Research, 103*(3), 409–420.
Rolls, E. T., Tovee, M. J., Purcell, D. G., Stewart, A. L., & Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates, and face identification and detection. *Experimental Brain Research, 101*(3), 473–484.
Rousselet, G. A., Macé, M. J. M., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision, 3*(6), 440–455 5.
Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research, 37*(3), 495–510.
Sato, T. (1989). Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Experimental Brain Research, 77*(1), 23–30.
Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (V1) activity gates awareness of motion. *Nature Neuroscience, 8*(2), 143–144.
Streit, M., Ioannides, A. A., Liu, L., Wölwer, W., Dammers, J., Gross, J., ... Müller-Gärtner, H. W. (1999). Neurophysiological correlates of the recognition of facial expressions of emotion as revealed by magnetoencephalography. *Cognitive Brain Research, 7*(4), 481–491.
Sweeny, T. D., D'Abreu, L. C., Elias, E., & Padama, L. (2017). Object-substitution masking weakens but does not eliminate shape interactions. *Attention, Perception, & Psychophysics,* 1–11.
Sweeny, T. D., Grabowecky, M., Kim, Y. J., & Suzuki, S. (2011). Internal curvature signal and noise in low-and high-level vision. *Journal of Neurophysiology, 105*(3), 1236–1257.
Sweeny, T. D., Grabowecky, M., Paller, K. A., & Suzuki, S. (2009). Within-hemifield perceptual averaging of facial expressions predicted by neural averaging. *Journal of*

Vision, 9(3), 1–11 2.

Sweeny, T. D., Suzuki, S., Grabowecky, M., & Paller, K. A. (2013). Detecting and categorizing fleeting emotions in faces. *Emotion, 13*(1), 76–191.

Suzuki, S., Clifford, C., & Rhodes, G. (2005). High-level pattern coding revealed by brief shape aftereffects. In C. W. Clifford, & G. Rhodes (Vol. Eds.), *Fitting the mind to the world: Adaptation and after-effects in high-level vision: vol. 2*, (pp. 135–172). Oxford University Press.

Tanaka, K., Saito, H. A., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology, 66*(1), 170–189.

Tomasik, D., Ruthruff, E., Allen, P. A., & Lien, M. C. (2009). Nonautomatic emotion perception in a dual-task situation. *Psychonomic Bulletin & Review, 16*(2), 282–288.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242–249.

Van Vugt, M. A. R. K., & Kameda, T. (2012). Evolution and groups. *Group Processes,* 297–332.

VanRullen, R., & Koch, C. (2003). Visual selective behavior can be triggered by a feed-forward process. *Journal of Cognitive Neuroscience, 15*(2), 209–217.

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30*(6), 655–668.

Webb, T. W., Igelström, K. M., Schurger, A., & Graziano, M. S. (2016). Cortical networks involved in visual awareness independent of visual attention. *Proceedings of the National Academy of Sciences, 113*(48), 13923–13928.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168.

Whitney, D., Haberman, J., & Sweeny, T. D. (2014). From textures to crowds: Multiple levels of summary statistical perception. In J. S. Werner, & L. M. Chalupa (Eds.). *The New Visual Neurosciences* (pp. 695–710). MIT Press.

Williams, M. A., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2005). Differential amygdala responses to happy and fearful facial expressions depend on selective attention. *Neuroimage, 24*(2), 417–425.

Woodman, G. F., & Luck, S. J. (2003). Dissociations among attention, perception, and awareness during object-substitution masking. *Psychological Science, 14*(6), 605–611.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256*(5061), 1327–1331.

Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *Journal of Neuroscience, 25*(36), 8150–8164.