HHS Public Access

Social Vision: Functional Forecasting and the Integration of Compound Social Cues

Reginald B. Adams Jr. and
Department of Psychology, The Pennsylvania State University, University Park, PA, USA

Kestutis Kveraga
Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Abstract

For decades the study of social perception was largely compartmentalized by type of social cue: race, gender, emotion, eye gaze, body language, facial expression etc. This was partly due to good scientific practice (e.g., controlling for extraneous variability), and partly due to assumptions that each type of social cue was functionally distinct from others. Herein, we present a functional forecast approach to understanding compound social cue processing that emphasizes the importance of shared social affordances across various cues (see too Adams, Franklin, Nelson, & Stevenson, 2010; Adams & Nelson, 2011; Weisbuch & Adams, 2012). We review the traditional theories of emotion and face processing that argued for dissociable and noninteracting pathways (e.g., for specific emotional expressions, gaze, identity cues), as well as more recent evidence for combinatorial processing of social cues. We argue here that early, and presumably reflexive, visual integration of such cues is necessary for adaptive behavioral responding to others. In support of this claim, we review contemporary work that reveals a flexible visual system, one that readily incorporates meaningful contextual influences in even *nonsocial* visual processing, thereby establishing the functional and neuroanatomical bases necessary for compound social cue integration. Finally, we explicate three likely mechanisms driving such integration. Together, this work implicates a role for cognitive penetrability in visual perceptual abilities that have often been (and in some cases still are) ascribed to direct encapsulated perceptual processes.

As human beings, we are reliant on each other for nearly all of our survival needs. Perhaps not surprisingly then, we have evolved an elaborate system of nonverbal exchange that allows us to effortlessly establish and maintain cohesive social interactions, to alert one another of potential rewards, and to warn against potential threats. If the ability to produce and recognize facial expressions evolved as a direct function of communicative utility, one argument is that such an adaptation should be less evident in environments where visual cues are obscured, as in the dark or in a forest environment where vision is obstructed by dense foliage.

Department of Psychology, Pennsylvania State University, 464 Moore Building, University Park, PA, 16802, RAdams@psu.edu.

Supporting such a supposition is work showing that terrestrial Old World monkeys compared with arboreal New World monkeys possess a greater repertoire of facial expressive potential (Redican, 1982). Further, it has been argued that as we evolved as a species, we became our own competitors for physical and social resources, necessitating a fine-tuned ability to mask true intentions and feelings, which in turn required even further refinement of our social perceptual abilities (Dunbar, 1998). Such a complex game of social chess presumably triggered a quantum leap in the evolution of our neocortical development, a contention supported by evidence that when comparing primate species (including humans), average neocortical volume correlates with average social network size (Dunbar).

Taken together, this work suggests that evolution has prepossessed us with the ability to communicate and understand a complex nonverbal language. Sapir put it well, stating that *"we respond to gestures…in accordance with an elaborate and secret code that is written nowhere, known by none and understood by all"* (1927, p. 892), suggesting that our ability to engage in nonverbal exchange is innately prepared. From a neurological perspective, this could be used to presume a "hardwired" highly modularized aspect of our biological make up that governs fixed perceptual functions. Recent evidence, however, has begun to question such assumptions, calling instead for a more flexible understanding of socio-emotional perception, and as such, of the organization of mind and brain function.

Before delving into this recent research, however, we review historical accounts of neural modularity underlying discrete and encapsulated processes thought to be involved in emotion and face processing and the assumptions these accounts gave rise to in our understanding of the functional neural architecture of the human brain. Then, we review recent research examining the strong contextual influences that are evident even for non-social visual processes, which offer dissenting evidence to previous accounts of direct perceptual encapsulation. Finally, through a conceptual merger between this new understanding of the perceptual mechanics of vision and a functional affordance account of compound social cue processing, we discuss how various facial and bodily cues unite in perception to yield the type of unified representations of others that govern our impressions of and behavior toward them.

The Case for Encapsulated Visual Perception

Modularity is often thought to be domain-specific (e.g., only responsive to visual input), inaccessible to conscious awareness or control, and encapsulated (unaffected by other types of input; e.g., visual processing should not be influenced by, for example, auditory, tactile, or olfactory cues, or by prior knowledge/expectations). An often-cited example of such a domain-specific, and cognitively impenetrable process, is the Müller-Lyer illusion (see also Fodor, 1983). This illusion, in which a line segment is enclosed either by two arrow points or two arrow ends produces the illusion that the line segment enclosed by arrow ends is longer, even though both line segments are of identical length (Fig. 1a). Likewise, in the Ebbinghaus illusion, a circle enclosed by smaller circles is perceived as larger than an identically sized circle enclosed by larger circles (Fig. 1b). Such misjudgments occur even with explicit knowledge of the illusions, and are exhibited by both humans and monkeys (Tudusciuc & Nieder, 2010).

The study of neural modularity gained scientific prominence early on with the work of German neurologist and anatomist Korbinian Brodmann in early 20th century. Brodmann (1909) believed that cortical regions had very specific functions and that this specificity had to be reflected by their histology. This turned out to be partly true in the early visual cortex, but for the most part, architectonic boundaries are insufficient for relating brain tissue to function, and therefore knowledge of other parameters – such as the regional connectivity patterns and neuronal tuning – are necessary to determine the function of a cortical area (Kaas, 2000).

Using vision as a model system, single-unit neurophysiology studies, starting with the seminal work of Hubel & Wiesel (1959), revealed a high degree of modularity and specialization, at least early in the processing hierarchy. Visual cortical areas such as V1, V2, V3, V4, V5 and so on were found to specialize in different "basic" aspects of visual stimuli, such as contours, color, form, motion and so forth. Then, with the advent of functional neuroimaging in the 1990s and 2000s, there were a flurry of brain mapping studies, which were markedly modular in their approach. Neatly, memorably labeled modules were proposed for the processing of higher-level stimuli (see Kanwisher, 2010), such as faces ("the fusiform face area", Kanwisher, McDermott, Chun, 1997), bodies ("the extrastriate body area"; Downing, Jiang, Shuman, Kanwisher, 2001), places ("the parahippocampal place area", Epstein & Kanwisher, 1998), and word forms (the "visual word form area"; Cohen, Dehane et al. 2000).

In sum, for many years the notion of neural modules supported scientists' assumptions that visual perception (from contours to faces) was the product of feed-forward processes that were considered, for the most part, cognitively impenetrable. That is, the perceptual experiences of different observers, given the exact same external circumstances, would be exactly the same and unaffected by influential cognitive states or other sensory modalities (Siegel, 2011). Although this effectively may be the case for certain very low-level visual perceptual experiences, we aim to show here that visual perception, and especially social visual perception, is deeply permeated by the observer's internal state, prior knowledge, and context, and as such, is highly cognitively penetrable.

A Case for Cognitive Penetration in Visual Perception

To understand what our senses are telling us, we must rely on some prior knowledge that links the sensations to some concept. For example, a particular patterning of the facial muscles and features in another's face could be interpreted as that person feeling angry, while the ocular muscles (eye gaze direction, pupil dilation) may provide additional information about the source and target of the anger. However, it is important to keep in mind that while such mappings can evoke conscious high-level concepts, they need not. Evolution has ensured that presumably non-conscious animals are able to recognize danger stimuli and mating signals from their conspecifics (see also Weisbuch & Adams, 2012). We will expand upon these points shortly, but first let us take a step back and examine these processes in *non*-social vision.

It has been known for quite some time that the process of vision involves more than just visual cortices, but also many association cortices in the temporal, parietal, and prefrontal lobes. Indeed, the tremendous brain expansion in hominids occurred primarily in the association cortex, not sensory or motor cortices (Buckner & Krienen, 2013). While the area of the primary visual cortex is similar in humans, macaques, and chimpanzees (accounting for brain size), it is the association cortices that contribute to the disproportionately large brain size in humans. Given that brain tissue is metabolically expensive, all this cortex must be doing something useful to promote our survival. Vision, our primary sense, involves a substantial portion of the association cortex, which is devoted to interpreting sensory input and generating predictions (Kveraga, Ghuman, Bar, 2007; Kveraga, Boshyan, & Bar, 2009). While much work remains to be done in resolving the exact functions of particular association regions, it is beginning to become apparent that chunks of tissue specialized for certain features act in a coordinated fashion to analyze afferent input, predict proximal events, and generate appropriate responses.

Consider a situation that occurs daily for many of us in the developed world - crossing a busy intersection. To do so without being hit, we must recognize a multitude of objects (cars, trucks, buses, bikes, pedestrians) moving at different speeds and from different directions, and predict their likely paths and intentions. Drawing on experience, one might base predictions on cues such as facial expression, body language, gender, ethnicity, type and make of the vehicle, traffic light patterns, even time of day, weather conditions, and geographical location. Our brains learn the statistical patterns associated with these variables through experience[1] and then use them automatically. These complex contextual inferences can be computed in our brains without giving a conscious thought to the process, sometimes even while talking or texting on a mobile phone, and most people repeatedly survive such crossings unscathed.

The Neural Mechanics of Cognitive Penetration

Let us now briefly examine the putative brain mechanisms for how we accomplish such a difficult task with apparent ease. The two-dimensional image on our retinae activates several different types of cells that form the major visual pathways, the magnocellular, parvocellular, and koniocellular pathways. Of these, the former two have been the most studied, and contribute predominantly to the dorsal and ventral visual streams, respectively (Goodale & Milner, 1992). The coarser, faster magnocellular (M) pathway triggers processing of meaningful stimuli (Chaumon, Kveraga, Barrett, Bar, 2013) in the orbitofrontal cortex (OFC), which provides modulating feedback to the slower processing of identity in the ventral visual stream (Kveraga, Boshyan, and Bar, 2007). The M-dominated dorsal stream play important roles in motion perception of, and guiding of attention and eye movements to salient objects. In our street-crossing example above, identification of vehicle makes and colors requires efficient scene segmentation and template matching, processes which take place in the occipital and ventral temporal association cortices in coordination with dorsal and prefrontal association regions. Navigation and perception of the spatial

---

[1]Please see this Slate article on how drivers of certain car makes are perceived: http://www.slate.com/articles/life/a_fine_whine/ 2013/07/bmw_drivers_and_cyclists_the_war_between_the_luxury_cars_and_bicycles.html)

qualities of the street scene, and the evaluation of its context engages regions such as the hippocampus and the parahippocampal, retrosplenial, and medial prefrontal cortices. Lastly, inference of social cues, such as the operator's age, gender, ethnicity, and mental state, engages a vast network of cortical and subcortical regions, including the amygdala, the fusiform face area (FFA), extrastriate body area (EBA), superior temporal sulcus (STS), temporoparietal junction (TPJ), and OFC.

The coordination of activity in these interacting networks is only beginning to be understood. A study by Kveraga, Boshyan, and Bar (2007) examined the contributions of magnocellular and parvocellular pathways to object recognition. As introduced above, recognition of objects must involve the invocation of memory templates that match – at least approximately – visual input. The question is, how are these object memories triggered, and by what mechanism? The authors employed line drawings of objects altered to bias visual processing towards either the M (coarse, fast, achromatic) or P (detailed, slower, color-sensitive) pathway. They found that stimuli biased toward the M pathway activated the orbitofrontal cortex, while P-based stimuli activated the object-sensitive regions in the ventral occipito-temporal cortex. Increased activation of OFC predicted faster recognition of M-biased stimuli (suggesting top-down facilitation), while increased activation of ventral temporal (fusiform) cortex predicted *slower* recognition of P-biased stimuli, suggesting more effortful processing in the ventral visual pathway for stimuli that do not benefit from facilitation from the M-dominant pathway. The subsequent information flow analysis using Dynamic Causal Modeling (Friston, 2003) showed increased feedforward flow from the occipital (low-level visual) cortex to the ventral object processing regions in the fusiform cortex for the P-biased stimuli. Conversely, M-biased stimuli induced stronger feedforward connectivity between the occipital cortex and OFC, and *feedback* connectivity between OFC and object-sensitive regions in the fusiform cortex (Kveraga et al., 2007). The findings from this study thus indicate biasing of perceptual object recognition processes in the ventral visual stream by "top-down" feedback triggered in OFC via the magnocellular pathway.

What triggers these OFC associations? Is it any M-biased, low spatial frequency visual input, or is it more specific than that? A recent study presented low spatial frequency (LSF) images of common objects and Gabor patches (Chaumon, Kveraga, Barrett & Bar, 2013). While both types of stimuli were biased towards magnocellular processing, only meaningful stimuli – objects – activated OFC. Gabor patches evoked activation only in low-level visual cortex. Moreover, the meaningfulness of object stimuli was positively correlated with OFC activation and its interactions with the fusiform object recognition regions (Chaumon et al., 2013). These findings again suggest that visual input matching stored memory associations triggers top-down facilitation processes in OFC, which exerts biasing feedback on task-specific analysis regions in the ventral visual pathway.

Another recent study using fMRI and MEG looked at a different type of top-down influence, investigating how contextual associations influence object perception using MEG and connectivity analyses (Kveraga et al., 2011). Objects were classified into two categories by prior rating studies – those having strong or weak links to some context (such as "kitchen", or "city street"). Images of strongly contextual objects, compared with weakly contextual objects, evoked phase-locking of activity between the early visual cortex and scene-sensitive

PHC at around ~150 ms, followed shortly by phase-locking between PHC and another scene-sensitive region RSC. RSC activity then was phase-locked with the early visual cortex at ~300 ms, which could be indicative of feedback processing from RSC to the visual cortex, and with medial prefrontal cortex at ~400 ms. PHC and RSC in particular have been known to be activated by scene images (Aguirre et al., 1996; Epstein and Kanwisher, 1998), but the stimuli in this study were objects that differed only in their contextual associations. The loci of these activations replicated findings in a series of studies (Bar and Aminoff, 2003; Aminoff et al., 2007; Bar, Aminoff, Schacter, 2008; Peters et al., 2009) in which stimulus context evoked activity in these regions.

These findings demonstrate the influence of context on what would be considered nonsocial stimuli – those depicting lines and geometric figures, objects, and scenes. That contextual factors can trigger powerful top-down modulation of such nonsocial visual perception (Kveraga et al., 2011) raises the important question about the extent to which contextual influences also act to guide social sensory perception. As will be seen in our review below, classic models of visual perception and emotion recognition, through years of assumptions of modularity and perceptual encapsulation, have fallen short of adequately considering the powerful role that context, prior knowledge, and motivation can have in social sensory perception.

Social Visual Perception

Some prominent face processing and emotion theories have speculated that there exist sub-modules implicated in the processing of distinct sources of facial information. One prominent model, for instance, contends that distinct sources of information derived from the face (e.g., identity versus expression) are processed along functionally dissociable and thereby noninteracting pathways (Bruce & Young, 1986; LeGal & Bruce, 2000). Research showing that the fusiform gyrus is implicated in the predominant processing of static appearance relevant aspects of the face, whereas the superior temporal sulcus (STS) is implicated in the predominant processing of the "changeable" aspects of the face, offers support for dissociable pathways (see Haxby et al., 2000), but this should not be taken to support *noninteractive* processing (see Townsend & Wenger, 2004 for theory of interactive parallel processing).

Likewise, within the domain of emotion perception, prevailing theories of emotion drew upon the seminal work of Darwin's *On the Expression of the Emotions in Man and Animals* (1872/1965; see also 1998, Paul Ekman's annotated version) to support the view that humans have evolved distinct and universal affect programs (i.e., discrete emotions such as anger, fear, happiness) that enable us to experience, express, and perceive emotions (Keltner, Ekman, Gonzaga, & Beer, 2003). Thus, it has been assumed that we possess evolved modules for adaptive responding and perceptual recognition of emotional behavior that have conferred upon us a distinct survival advantage. The obligatory nature of this process, as a result, for years left little room for considering how contextual influences, knowledge, and individual states may impact emotion perception.

Recently, however, vision scientists have become increasingly cognizant of the powerful role context plays in vision, particularly in social vision (see Adams, Ambady, Nakayama, & Shimojo, 2010). In order for us to function and succeed as social beings, a primary factor guiding vision, and the evolution of the extended neocortical system, must include the people who fill our social worlds. Thus to fully understand the visual system in all its complexity, we must consider the socially adaptive functions it evolved to perform. Social visual perception is in its very essence a predictive process in which both stored memories and innate templates play a critical role. To be successful at visual perception, therefore, we must recognize even subtle visual cues—the tension in the jaw, flushed skin, contracted pupils—and interpret them quickly and correctly. Recognizing something by definition requires knowledge of the thing that is being recognized. Therefore, over the course of our evolution and development we would have had to acquire the mapping of sensory stimulation (e.g., on the retinae) to concepts such as "this person is angry towards me."

Despite this, descending from early assumptions of functional and neural modularity, social visual cues such as race, gender, age, appearance, eye gaze, and emotion were for decades studied in separate fields of inquiry, with other social cues treated as noise to be controlled. Emotional expression, for example, had been most extensively studied within the domain of social psychology, and facial identity largely within the domain of visual cognition. Eye gaze has been extensively studied within both fields, but with emphasis on social communicative value in social psychology and on perceptual mechanics in visual cognition (see also Adams & Nelson, 2012). Consequently, neither discipline adequately addressed the mutual contextual influences various social cues have on each other in visual perception.

As evidence has been accumulating in support of functional interactions across an ever-widening array of social cues, the need for new conceptual approaches is becoming evident. Initial evidence for functional interactivity was demonstrated through differential behavioral and neural responses to various combinations gaze direction and emotion expression (Adams & Kleck, 2003; Adams et al., 2003; see also Adams et al., 2012). A recent connectionist model likewise focuses on perceptual interactions of identity cues in the face (e.g., sex, race, age; Freeman & Ambady, 2010). Still, little is understood regarding the visual dynamics driving these interactions, particularly across the much broader set of cues available to our visual system. For instance, numerous studies now reveal functional interactions across a wide range of social cues including (to name a few): 1) gaze and head postures (Chiao, Adams, Tse, Lowenthal, Richeson, & Ambady, 2008), 2) race and emotion (Ackerman et al., 2006; Hugenberg, 2005; Weisbuch & Ambady, 2010), 3) race and eye gaze (Adams, Pauker, & Weisbuch, 2010; Richeson, Todd, Trawalter, & Baird, 2008; Trawalter, Todd, Baird, & Richeson, 2008), 4) gender and emotion (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Hess, Adams, & Kleck, 2004; 2005; cf. LeGal & Bruce, 2002), 5) facial expression and emotional body language, 6) facial expression and approach/avoidance movement (Adams, Ambady, Macrae, & Kleck, 2006), and 7) gaze and facial attraction (Jones et al., 2010). It is clear that such a complex array of social cues must combine at some point to provide information about another person's internal states and intentions (see Baron-Cohen, 1995). To examine when and where along the perceptual stream social cues are combined in perceptual experience, Adams and colleagues have proposed a functional forecasting account of this combinatorial process (Adams et al., 2010; Adams & Nelson, 2011;

Weisbuch & Adams, 2012). This approach draws heavily from existing ecological models of vision and social perception (Gibson, 1979; Zebrowitz, 1997). It is, therefore, critical to consider Gibson's original notions of perceptual affordances and attunements in visual perception.

Gibson defined affordances as opportunities to act on or be acted upon that are inherent in a stimulus. Thus when one sees a plate of food or glass of water, part of that perceptual experience is linked to the functional behavior associated with those objects, i.e., eating and drinking, respectively. Gibson argued that such affordances are processed via direct perception, i.e., not requiring the need for abstract cognitive intervention (and thus no cognitive penetration). Though direct in this sense presumably means evoked automatically, it is critical to point out that such affordances nonetheless can be and often are social in nature (Zebrowitz et al., 2010). When seeing an infant, for instance, caretaking behaviors of holding and protecting may be triggered automatically, provided that the perceiver is attuned to them. Gibson defined attunements as a perceiver's sensitivity to the affordances inherent in a stimulus, which can be shaped by evolution, but also can be the product of individual differences and situational contexts. Thus new parents might be expected to be more attuned to the sight and sounds of infants, particularly their own. Indeed, mothers have been shown to respond with heightened skin conductance responses (Hildebrandt & Fitzgerald, 1981; Wiesenfeld & Klorman, 1978) and cardiac acceleration (Wiesenfeld & Klorman, 1978) when presented with pictures of their own versus unfamiliar infants newborns. As such, Gibson emphasized that vision is shaped by the interaction between the perceiver and the perceived. From this perspective then it should not be too surprising that there are top-down influences evident in the recognition of even common objects (i.e., Kveraga, Boshyan, & Bar, 2009).

The relevance all this has to compound social cue processing becomes clear when we begin to examine the social meanings that are shared by various social cues. If various social cues conspire to signal the same behavioral affordance, we should then expect those cues to become perceptually integrated in our visual experience as well, and in a relatively reflexive manner. Such integration implies that in combination these cues mutually signal same underlying adaptive value or generate an entirely new signal value that is dependent on their mutual influence.

In terms of shared signals, of particularly relevance here are the motivational stances of dominance and affiliation, which are thought to be key organizing factors in social behavior and perception (Wiggins & Broughton, 1991; Schaefer & Plutchik, 1966). The perception of dominance and affiliation can be derived from many social cues, including movement of head, hands, and legs (Gifford, 1991), facial expressions (Knutson, 1996), facial appearance (Zebrowitz, 1997), even vocal cues (Ambady et al., 2002). In this way, the human form is replete with otherwise distinct social cues that nonetheless share ecologically relevant meaning. Although little work has examined the combinatorial nature of dominance and affiliation perception at the neural level in humans, neurophysiological studies in nonhuman primates reveal distinct neural mechanisms for their perception (Pineda et al., 1994). Underlying these basic social motives are even more fundamental tendencies driving biological movement, i.e., approach/avoidance (Baron-Cohen, 1995; Davidson & Hugdahl,

1995), which also appear to have a distinct neural basis in nonhuman primates (Brothers & Ring, 1993). Given the shared signal value associated with various social cues conveyed by the face and body, it makes adaptive sense then that combinations of these cues would be perceptually integrated to forecast the basic tendencies to approach, avoid, dominate, and affiliate. Such combinatorial processing arguably would be critical for ensuring a timely assessment of, and reaction to, social cues conveyed by others. Cues may also forecast information that is only meaningful in their combined association. For instance, if someone makes a fearful expression, they are signaling danger somewhere in the environment. Only in combination with gaze then does this expression forecast what direction you should run for safety.

This functional approach to compound social cue processing, therefore, focuses on the underlying meaning conveyed by various cues and their combined ecological relevance to the observer. We argue that because of their shared functional affordances, only in combination do social cues truly inform an adaptive behavioral response in an observer. Below, we examine more closely the perceptual dynamics of such compound cue integration with emphasis on three likely drivers including: stimulus-bound integration, feedforward, and feedback processing. Because some cues share physical resemblance, and thus are stimulus-bound, we do not expect integration of these cues to involve cognitive penetration. For feedforward integration the question is more difficult one to decipher. Feedback processing, of the kind that informs our conscious percepts, however, more clearly implicates the type of cognitive penetration that we have already reviewed for non-social visual processes.

A Case for Stimulus-Driven Integration

In Gibson's ecological model of vision (1979), the proposal that vision is bound to action suggests that important insights regarding the underlying mechanics of the perceptual system might be gleaned from deep examination of the stimulus itself. Gender-related facial appearance, emotion, and facial maturity offer good examples of facial cues that not only convey similar messages (dominance and affiliation), but do so in part through confounded facial cues. Anger is characterized by a low and bulging brow ridge and narrowed eyes, resembling a mature, dominant face, while fear with its raised and arched brow ridge and widened eyes perceptually resembles a submissive "babyface". Gender appearance is similarly associated with facial features that perceptually overlap with facial maturity (see Zebrowitz, 1997), in that "babyish" features (e.g., large, round eyes, high eye brows, full lips) are more typical in women, and "mature" features (e.g., square jaw, pronounced and lower brow ridge) are more typical in men. These gender cues similarly overlap with expression, with masculine faces physically resembling anger, and feminine faces physically resembling aspects of both fear and happiness (Hess, Adams, Grammer, & Kleck, 2009). Insofar as form follows function in this regard, the confounded nature of gender, emotion, and facial maturity as social cues to dominance and affiliation may shed important insight into how our perceptual system is functionally organized to integrate compound cues that mutually convey shared social affordances.

Recent theory resonates with these ideas suggesting that facial expression may have evolved to mimic stable facial appearance cues in order to exploit their social affordances (Adams et al., 2010; Becker et al., 2007; Hess et al., 2009; Marsh, Adams, & Kleck, 2005). This proposition echoes Darwin's early observation of piloerection in animals, which makes them appear larger and thus more threatening to ward off attack (see for example, 1865; p. 95 and p.104). Such an account offers a reasonable explanation for the existence of confounded facial cues across gender-related appearance and expression. Critically, according to this account compound cue integration necessarily begins at the level of the stimulus itself.

Another case in point descends from observations of comorbidity in deficits found in certain disorders affecting face processing. Most cases of prosopagnosia, for example, show some evidence for impaired emotion processing as well (Calder & Young, 2005). If these perceptual processes are functionally separable, however, we should expect more cases with clear, doubly dissociable deficits. To address such overlap, one study employed a principal component analysis (PCA) to serve as a statistical analogue for a system that might integrate appearance and expressive cues in a face (Calder, Burton, Miller, Young, & Akamatsu, 2001). Notably, PCA models such as this have previously been used to simulate a variety of face perception phenomena, including cross-race memory (O'Toole et al., 1993). Although this model provided evidence adequate to model the independent contributions of identity and expression, there was also a measurable degree of overlap between them. Because PCA is based on facial metric analysis, this finding emphasizes again the integration of identity and expressive cues beginning at the level of the stimulus itself, before afferent stimulation even hits the sensory system.

"Feed-forward" Integration?

Recently, de Gelder and Tamietto (2010) reviewed the evidence for the existence of very early social vision, in which processing may begin already in subcortical structures. As evidence, they highlight work showing that body language, visual scenes, and vocal cues all influence the processing facial displays of threat (Meeren, Hvan, Heijnsbergen, & de Gelder, 2005), even at the earliest stages of face processing and across conscious and nonconscious processing routes (de Gelder, Morris, & Dolan, 2005). Such findings, they argue, suggest that at least some compound cue integration is due to bottom-up influences governed by the shared functional value of these cues (in this case all expressing the same emotion). If true, this suggests that the visual system may be calibrated to respond to shared social affordances of compound social cues at the very earliest stages of perceptual processing, likely yielding pre-attentive and functionally diffuse (i.e., triggering appetitive and defensive motivational systems) responses (see also Weisbuch & Adams, 2012).

Compound cue integration that occurs in response to subliminally presented stimuli offers preliminary evidence for such nonconscious integration. One such study involved functional interactions in processing various pairings of eye gaze direction (direct versus averted) and threat displays (anger versus fear). In line with the *shared signal hypothesis* (Adams, Ambady, Macrae, & Kleck, 2006; Adams & Kleck, 2003, 2005), cues relevant to threat that share a congruent underlying signal value were predicted to facilitate the processing efficiency of an emotion. In support of this hypothesis, using speeded reaction time tasks

and self-reported perception of emotional intensity, Adams and Kleck (2003, 2005) found that direct gaze facilitated processing efficiency, accuracy, and increased the perceived intensity of facially communicated approach-oriented emotions (e.g., anger and joy), whereas averted gaze facilitated processing efficiency, accuracy, and perceived intensity of facially communicated avoidance-oriented emotions (e.g., fear and sadness). Adams et al. (2003), using fMRI, went on to find evidence for interactivity in amygdalar responses to these same threat-gaze pairings, but in this study, amygdala responses were found to be greater to threat-related ambiguity as compared to threat-congruent combinations. Notably, these effects were found when presenting stimuli for very sustained durations (2 seconds).

Later, Mogg, Garner, and Bradley (2007) speculated that the above findings may have been at least in part due to self-regulatory diversion of attention away from clear threat-gaze combinations, leaving more sustained attention to decipher threat-related ambiguity. They suggested that very salient/congruent cues may be more likely to influence reflexive threat responses: "the amygdala may indeed modulate attention to threat, but its level of activation may be a function of both initial orienting and maintained attention" (p. 167). In order to examine this possibility, we examined responses to subliminal (i.e., ~33 ms backward-masked) presentations of threat-gaze pairs, and in this case found greater right amygdala activation, as well as overall greater activation in extended visual, cognitive, and motor processing networks, but critically this time in response to congruent versus ambiguous threat cues. Notably, interdependencies between eye gaze and emotion such as this can not be attributable to the type of visually confounded properties we discussed in the last section, as these two facial cues occupy nonoverlapping space in the face (i.e., eyes can change direction without influencing facial muscle patterning and vice versa). Thus it would appear that low-level functional integration in this case is due to the functional affordances shared by these facial cues even when both are processed outside of conscious awareness.

Using a similar logic, Tamietto & De Gelder (2010) describe a series of studies in which patients with visual blindsight (i.e., with striate cortex damage that impairs conscious visual perception in either the left or right visual field) were asked to respond to a facial expression presented in their unimpaired visual field. In one set of studies, they found that recognition of expressions in the unimpaired visual field was speeded up if a stimulus person making an identical expression was presented in the "blind" visual field (de Gelder et al., 2005; de Gelder et al., 2001). Thus, information from the expression presented in the "blind" visual field appeared to directly influence, presumably via subcortical projections to extrastriate visual areas bypassing V1, the expression presented in the contralaterally presented unimpaired visual field. Even more compelling was evidence that responses to facial expressions presented in the unimpaired visual field were speeded when the same emotion was expressed by a body in the "blind" visual field (Tamietto, Weiskrantz et al., 2007). This finding suggests that the bodily expression of emotion, even though nonconscious, is integrated in the response to the facial display of emotion, despite such processing involving highly divergent visual properties. Thus, the key to integration in these studies appears once again to be due to shared social affordances. Finally, in a related study, Ghuman et al. (2010) offered additional evidence for cross-cue integration by demonstrating face adaptation effects in response to viewing bodies. In their study, they employed visual adaptation to human figures (displayed without a face) and showed that adapting to male or female shapes

shifted the judgments of face gender as well. This study thus demonstrated that viewing a human body directly shapes the perceptual experience of a visually dissimilar, even cross-category stimulus – in this case the face.

The notion that certain social vision functions may be carried out starting in early subcortical structures is quite provocative. Evolutionarily old subcortical structures, such as the superior colliculus (SC), for instance, do appear to be sensitive to threat cues (Maior et al. 2011). SC (known as the optic tectum in non-mammals) has projections to the koniocellular (K) layers of the lateral geniculate nucleus, and also projects to the pulvinar and mediodorsal nuclei of the thalamus. The latter thalamic nuclei project to many cortical and subcortical regions involved in social and threat perception, such as the amygdala. Though no monosynaptic or polysynaptic pathway has been yet identified, these connections may constitute a non-geniculocortical pathway similar to the "low road" for rapid threat detection in audition (LeDoux, 1996). This putative pathway has been proposed to be responsible for right amygdala responses to subliminally presented threat stimuli, out of participants' conscious visual experiences (Morris, Öhman, & Dolan, 1998; see also Öhman, 2005), though again this has yet to be demonstrated by anatomical tracing studies. Therefore, it is debatable whether this pathway is or even needs to be entirely subcortical or feedforward in nature (Pessoa & Adolphs, 2010). The most likely scenario is that several different pathways converge on the structures relevant in social visual perception, supplying different types of information and influence. This may include luminance-based contrast information (from M, P and K pathways), chromatic information of different wavelengths (P and K pathways), corollary discharges conveying information about internal state, current goals, eye position, and the like.

What is clear is that some pathways that bypass the primary visual cortex (and thus, conscious awareness) do appear to exist, and are particularly responsive to low spatial frequencies, implicating magnocellular system involvement (Vuilleumier, Armory, Driver, & Dolan, 2003) and/or the evolutionarily ancient and likewise low spatial frequency koniocellular pathway involvement (Hendry & Reid, 2000; Morand et al., 2000; Xu et al., 2001; Isbell, 2006; Jayakumar et al., 2013). These pathways feed heavily into the dorsal visual stream via projections to area MT, a major waystation to action-related vision, which is a neural underpinning necessary to support Gibson's notion of affordance-based vision (see also Nakayama, 2010). Gibson's notion of direct perception, however, implies that this system does not require cognitive intervention. In the next section, we revisit visual illusions to help dissociate "action" vision from later conscious perceptual experiences, which suggest that at least some magnocellular visual experiences do involve direct perception that can be dissociated from later cognitively penetrated visual experiences.

Revisiting the Case for "Top-down" Modulation in Visual Experience

As reviewed above, there now exists an extensive examination of the impact of context on even basic object recognition (see Kveraga, Boshyan, & Bar, 2007; Kveraga, Ghuman, & Bar, 2007; Kveraga, Boshyan, and Bar, 2009). Given that the neural mechanics of vision support such top-down contextual influence, we should expect even more associative influences to be apparent in social perception. Thus, even if compound social cues are not

integrated in a bottom-up or feedforward fashion, due to overlapping physical properties and/or shared neural substrates as proposed above, we know that top-down modulation of even low-level visual processing occurs quite rapidly, thereby potentially impacting very early integration. For instance, as previously reviewed, magnocellular projections to association areas in the brain (e.g., orbitofrontal cortex) quickly project back to visual centers, helping guide even simple object recognition (Kveraga, Boshyan, & Bar, 2007), and all within the first ~200 ms of visual input (Bar et al., 2006). Thus, we must consider the importance of top-down integration of contextual meaning when guiding and organizing social visual processing. In this way, compound cues can still be quickly integrated in a functionally meaningful way.

Let us consider for a moment another visual illusion that relies on top-down expectations in face processing, the "hollow-mask" illusion. In this illusion, even proving to the observers that they are viewing the concave part of a mask, by rotating the mask 360°, does not dissipate the conscious perception of viewing a convex, 3D face. Such an illusion might therefore be considered an example of cognitive impenetrability in the sense that it appears to resist willful efforts to correct for the top-down effects – in that perceivers struggle to see the concave side of the mask for what it is, but rather perceive it as a "normal" (i.e., 3D, convex) face. What scientists believe this illusion actually demonstrates, however, is very strong evidence for the powerful influence of top-down expectations on an otherwise very low-level perceptual experience (Gregory, 1970). Emphasizing this point is work now further demonstrating that there exists a dissociation between the M-dominant, reflexive visual system, which triggers accurate behavioral responses to the hollow mask (participants could flick the mask with precision suggesting something akin to Gibson's proposition for direct action-related vision), and the P-dominant, reflective system, which yields the conscious, and seemingly uncontrollable, experience of the erroneous concave percept (Króliczak et al., 2006).

For humans, other human faces are arguably the most ubiquitous and informative social stimuli we encounter in our daily lives. We see faces everywhere, even where none exist, such as in cloud formations, rock outcroppings, surfaces of other planetary objects (e.g., the Face on Mars, the Man in the Moon), even in pieces of burnt toast, through top-down expectations related to the powerful social significance of the human face. A recent MEG study underscores this point by investigating a type of pareidolia (i.e., seeing faces in common objects) finding that they indeed evoked neuromagnetic activity similar in timing and amplitude to real faces in the face-sensitive cortex in the ventral temporal lobe (Hadjikhani, Kveraga et al., 2009).

Further supporting a top-down influence in face processing are findings that social category membership can shape perceptual experiences. For instance, same- versus other-race faces are processed more holistically (i.e., encoding features in relation to one another such that the whole is greater than the parts), whereas other-race faces are processed in a more piecemeal manner (i.e., encoding features independent of each other) (see McKone, Brewer, MacPherson, Rhodes, & Hayward, 2007; Tanaka, Kiefer, & Bukach, 2004; Turk, Handy, & Gazzaniga, 2005). Consistent with these findings, participants show better memory performance for same-race faces, while also showing better memory for the individual

features of other-race faces (Michel et al., 2006). Given the role of holistic processing in face memory, it is perhaps not surprising then that even ambiguous-race faces when coupled with stereotypical hairstyles (Black or Latino) give rise to the own-race memory bias in memory (Maclin & Malpass, 2001). Notably, when coupled with Black versus Latino hairstyles these same faces were also rated as having darker skin tone, wider faces, and less protruding and wider eyes, further underscoring the top-down effects of conscious perceptual experiences. Pauker Weisbuch, Ambady, Sommers, Adams, & Ivcevic (2009) likewise found that racial labels assigned to racially ambiguous faces gave rise to own-race memory biases. Rule, Ambady, Adams, and Macrae (2007) similarly found heterosexual and homosexual participants showed an own-group face memory bias based on their perceptions of sexual orientation in target faces. Bernstein, Young, & Hugenberg (2007) found similar effects for arbitrarily assigned group memberships (i.e., same university or same personality type), as did Shriver, Young, Hugenberg, Bernstein, & Lanter (2008) based on the perceived social class of the stimulus faces. Thus it appears that when enough category-specific information is supplied to identify an ambiguous face as an outgroup member, it is then processed in a qualitatively different manner than that of an ingroup member.

In terms of compound social cue processing, race also appears to interact with other social cues in a combinatorial manner to influence memory. For instance, Adams, Pauker, and Weisbuch (2010) replicated the own race bias in memory for faces posed with direct gaze, but not averted gaze, and the effect was driven by better performance for same-race faces as opposed to worse performance to other-race faces, presumably again due to holistic face processing. This latter conclusion is consistent with evidence from Turk, Handy, and Gazzinaga (2005) who similarly found that in a split-brain patient, the own-race bias memory effects were only evident in the right hemisphere, which is known to underlie holistic face processing, again and effect due to enhanced memory for ingroup faces. Two additional studies examining the combined influences of race and gaze found that viewing Black relative to White faces generated more amygdala responsivity and selectively captured attention to a greater degree in White observers (Richeson, Todd, Trawalter, & Baird, 2008; Trawalter, Todd, Baird, & Richeson, 2008). This effect too was evident only for faces coupled with direct relative to averted gaze. They argued direct gaze heightened the threat-value of viewing Black faces (see also Adams & Kleck, 2003, 2005). Also related is research demonstrating that White participants' memory for Black versus White faces increases when coupled with angry versus neutral displays, an effect that was argued to be due to the functional value of perceiving compound threat cues (Ackerman et al., 2006). Both of these latter findings are highly consistent with the functional account proposed herein.

Race has also been found to influence how eye gaze is processed. That others' gaze direction cues observers' attention has been long thought to be an innately prepared and reflexive visual response (Frischen, Bayliss, & Tipper, 2007). Yet, even in monkeys, high-status members of a colony only exhibit such gaze following in response to other high-status monkeys, not low status monkeys (Shepherd, Deaner, & Platt, 2006), an effect replicated in humans when viewing masculine versus feminine faces (Jones et al., 2009). These findings suggest that in humans gaze-mediated visual attention might also be more prevalent when viewing models who hold privileged status. Consistent with this, Weisbuch, Pauker, Lamer,

and Adams (*under review*) recently found that White participants, who historically have held a privileged status in the U.S., only oriented attention to the gaze of White, not Black stimulus models, whereas Black participants oriented to the gaze of both groups. What is compelling about this finding is that a social group membership impacted what has long been considered an obligatory visual response, one not expected to be sensitive to top-down modulation.

Although to our knowledge, no studies to date have extended the existing work examining M-pathway contributions to object recognition to compound social cues processing, there is little reason to assume such associative processes would not contribute to social perception. If anything, we should expect M-pathway contributions to social vision to be even more powerful. One recent study, for instance, examining gender perception in the face, found that although responses in the fusiform cortex tracked closely with the objective linear gradations in face gender between male and female faces (manipulated using a morphing algorithm to average the texture and structural maps of male and female faces), OFC (which receives strong M-pathway projections) tracked with subjective (categorical, and thus nonlinear) perceptions of face gender (Freeman, Rule, Adams, & Ambady, 2009). This study reflects the important role of OFC in generating associative predictions about the nature of a social visual stimulus presented, and as such strongly underscores this as a fruitful, and necessary, direction for future social visual inquiry.

Conclusion

In this paper, we reviewed classic and contemporary literatures that speak to the issue of cognitive penetrability in visual perception, and examined in detail the neural mechanics that are necessary to allow for a "top-down" modulated guidance system to social visual perception. Having established that the neural dynamics exist to support contextual influences in nonsocial visual experience, we propose that social visual experience should be expected to involve even greater cognitive involvement, particularly with compound social cues from face and body being processed in a functionally relevant manner.

We proposed three mechanisms for such combinatorial processing. The first of these involves purely bottom-up integration of social cues that begins at the level of the stimulus itself. In this case the visual properties across different facial cues that specify similar social affordances are physically confounded, and thus are *not* likely implicating cognitive penetration. We then discussed potential nonconscious integration of even visually divergent cues that share social affordances, such as eye gaze, body language, and facial expression. Critically, that this integration may occur as part of a fast feed-forward sweep through the visual and emotion areas in the brain does not exclude top-down influence. Except for the simplest stimuli, even "bottom-up" perceptual processes always ultimately involve mapping afferent information onto known (and thus "top-down") associations to result in meaningful perception. Such integration likely occurs in oscillating networks of brain regions relevant for a particular task or thought process, with oscillation synchrony possibly signaling successful realization of this process (for a review, see Buszaki, 2006; 'Rhythms of the Brain"). Finally, we discussed the more clear role of top-down influences on social visual processing, particularly with regard to shared associations across cues that are of ecological

relevance to the observer reviewing effects of gaze perception and own-race biases. Critically, all three of these proposed mechanisms converge in terms of their central reliance on shared functional dynamics guiding compound cue integration, although each represents a vastly different mechanism for such integration. We speculate that all of these mechanisms are at play during social visual perception of both evolved and socially learned affordances.

The myriad of social information we forecast to one another, in the blink of an eye, results in a holistic impression that guides our adaptive behavioral reactions to one another. So, if a large figure with clenched fists and a hypermasculine visage (low brow, thin lips, angular) approaches you from an empty alley while staring directly at you and making an anger expression…RUN! But, also take note that you will likely perceive that facial expression of anger more clearly and efficiently than if you were to see the same expression worn by a child with a "babyish" face (high brow, full lips, round), in a crowd, looking away from you. In this way, the functional forecast framework we present here challenges strict assumptions of functional and neural modularity, and offers evidence for cognitive penetrability even at the level automatic and presumably reflexive visual experiences.

References

Ackerman JM, Shapiro JR, Neuberg SL, Kenrick DT, Becker DV, Griskevicius V, et al. They all look the same to me (unless they're angry): From out-group homogeneity to out-group heterogeneity. Psychological Science. 2006; 17:836–840. [PubMed: 17100781]

Adams, RB., JrAmbady, N.Nakayama, K., Shimojo, S., editors. The Science of Social Vision. New York, NY: Oxford University Press; 2010.

Adams, RB, Jr. Facing a perceptual crossroads: mixed messages and shared meanings in social perception. Talk presented at the annual meeting for Society for Personality and Social Psychology; Tampa, FL. 2009.

Adams RB Jr, Ambady N, Macrae CN, Kleck RE. Emotional expressions forecast approach-avoidance behavior. Motivation and Emotion. 2006; 30:177–186.

Adams, RB., Jr, Franklin, RG., Jr, Nelson, AJ., Stevenson, MT. Compound social cues in face processing. In: Adams, RB., JrAmbady, N.Nakayama, K., Shimojo, S., editors. The Science of Social Vision. Oxford University Press; 2010. p. 90-107.

Adams RB Jr, Franklin RG Jr. Influence of emotional expression on the processing of gaze direction. Motivation & Emotion. 2009; 33:106–112.

Adams RB Jr, Gordon HL, Baird AA, Ambady N, Kleck RE. Effects of gaze on amygdala sensitivity to anger and fear faces. Science. 2003; 300:1536. [PubMed: 12791983]

Adams RB Jr, Kleck RE. Perceived gaze direction and the processing of facial displays of emotion. Psychological Science. 2003; 14:644–647. [PubMed: 14629700]

Adams RB Jr, Kleck RE. Effects of direct and averted gaze on the perception of facially communicated emotion. Emotion. 2005; 5:3–11. [PubMed: 15755215]

Adams, RB., Jr, Nelson, AJ. Intersecting identities and expressions: The compound nature of social perception. In: Decety, J., Cacioppo, J., editors. The Handbook of Social Neuroscience. Oxford University Press; 2011. p. 394-403.

Adams RB Jr, Pauker K, Weisbuch M. Looking the other way: The role of gaze direction in the cross-race memory effect. Journal of Experimental Social Psychology. 2010; 46:478–481. [PubMed: 20204166]

Aguirre GK, Detre JA, Alsop DC, D'Esposito M. The parahippocampus subserves topographical learning in man. Cereb Cortex. 1996; 6:823–829. [PubMed: 8922339]

Ambady N, LaPlante D, Nguyen T, Rosenthal R, Levinson W. Surgeon's tone of voice: A clue to malpractice history. Surgery. 2002; 132:5–9. [PubMed: 12110787]

Aminoff E, Gronau N, Bar M. The parahippocampal cortex mediates spatial and nonspatial associations. Cereb Cortex. 2007; 17:1493–1503. [PubMed: 16990438]

Bar M, Aminoff E. Cortical analysis of visual context. Neuron. 2003; 38:347–358. [PubMed: 12718867]

Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hamalainen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E. Top-down facilitation of visual recognition. Proc Natl Acad Sci USA. 2006; 103:449–454. [PubMed: 16407167]

Baron-Cohen, S. Theory of mind and face-processing: How do they interact in development and psychopathology. In: Cicchetti, D., Cohen, D., editors. Developmental Psychopathology. Vol. 1. New York, NY: John Wiley & Sons; 1995. p. 343-356.

Becker DV, Kenrick DT, Neuberg SL, Blackwell KC, Smith DM. The confounded nature of angry men and happy women. Journal of Personality and Social Psychology. 2007; 92:179–190. [PubMed: 17279844]

Bernstein MJ, Young SG, Hugenberg K. The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. Psychological Science. 2007; 18:706–712. [PubMed: 17680942]

Brodmann, K. Localization in the cerebral cortex. Springer; 1909/2006.

Brothers L, Ring B. Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. Behavioural Brain Research. 1993; 57:53–61. [PubMed: 8292255]

Bruce V, Young A. Understanding face recognition. British Journal of Psychology. 1986; 77:305–327. [PubMed: 3756376]

Buckner R, Krienen F. The evolution of distributed association networks in the human brain. Trends in Cognitive Sciences. 2013; 17:648–665. [PubMed: 24210963]

Calder AJ, Young AW. Understanding the recognition of facial identity and facial expression. Nature Reviews Neuroscience. 2005; 6:641–651. [PubMed: 16062171]

Calder AJ, Burton AM, Miller P, Young AW, Akamatsu S. A principal component analysis of facial expressions. Vision Research. 2001; 41:1179–1208. [PubMed: 11292507]

Chaumon M, Kveraga K, Barrett LF, Bar M. Visual predictions in the orbitofrontal cortex rely on associative content. Cerebral Cortex. 2013; doi: 10.1093/cercor/bht146

Chiao JY, Adams RB Jr, Tse PU, Lowenthal WT, Richeson JA, Ambady N. Knowing who's boss: fMRI and ERP investigations of social dominance perception. Group Processes & Intergroup Relations. 2008; 11:201–214. [PubMed: 19893753]

Cohen L, et al. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. Brain. 2002; 125:1054–1069. [PubMed: 11960895]

Darwin, C. The expression of the emotions in man and animals. Chicago: University of Chicago Press; 1872; 1965; 1998.

Davidson, RJ., Hugdahl, K., editors. Brain asymmetry. Cambridge, MA: Mit Press; 1995.

de Gelder B, Morris JS, Dolan RJ. Unconscious fear influences emotional awareness of faces and voices. Proceedings of the National Academy of Sciences. 2005; 102:18682–18687.

de Gelder B, Pourtois G, van Raamsdonk M, Vroomen J, Weiskrantz L. Unseen stimuli modulate conscious visual experience: evidence from inter-hemispheric summation. Neuroreport. 2001; 12:385–391. [PubMed: 11209955]

de Gelder, B., Tamietto, M. Faces, bodies, social vision as agent vision and social consciousness. In: Adams, RB.Ambady, N.Nakayama, K., Shimojo, S., editors. The science of social vision. New York: Oxford University Press; 2010. p. 51-74.

Downing PE, Jiang Y, Shuman M, Kanwisher N. A cortical area selective for visual processing of the human body. Science. 2001; 293:2470–2473. [PubMed: 11577239]

Dunbar RIM. The social brain hypothesis. Evolutionary Anthropology. 1998; 6:178–190.

Epstein R, Kanwisher N. A cortical representation of the local visual environment. Nature. 1998; 392:598–601. [PubMed: 9560155]

Fodor, JA. The modularity of mind. Cambridge, MA: MIT Press; 1983.

Freeman JB, Rule NO, Adams RB Jr, Ambady N. The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. Cerebral Cortex. 2010; 20:1314–1322. [PubMed: 19767310]

Frischen A, Bayliss AP, Tipper SP. Gaze cueing of attention: Visual attention, social cognition, and individual differences. Psychological Bulletin. 2007; 133:694–724. [PubMed: 17592962]

Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage. 2003; 19:1273–1302. [PubMed: 12948688]

Gibson, JJ. The ecological approach to visual perception. Boston: Houghton-Mifflin; 1979.

Goodale MA, Milner AD. Separate visual pathways for perception and action. Trends in Neuroscience. 1992; 15(1):20–25.

Hadjikhani N, Kveraga K, Naik P, Ahlfors SP. Early (M170) activation of face-specific cortex by face-like objects. Neuroreport. 2009 Mar 4; 20(4):403–7. [PubMed: 19218867]

Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends in Cognitive Sciences. 2000; 4:223–233. [PubMed: 10827445]

Hendry SH, Reid C. The koniocellular pathway in primate vision. Annu Rev Neurosci. 2000; 23:127–153. [PubMed: 10845061]

Hess U, Adams RB Jr, Grammer K, Kleck RE. Sex and emotion expression: Are angry women more like men? Journal of Vision. 2009; 9:1–8.

Hess U, Adams RB Jr, Kleck R. Facial appearance, gender, and emotion expression. Emotion. 2004; 4:378–388. [PubMed: 15571436]

Hess U, Adams RB Jr, Kleck R. Who may frown and who should smile? Dominance, affiliation, and the display of happiness and anger. Cognition & Emotion. 2005; 19:515–536.

Hildebrandt KA, Fitzgerald HE. Mothers' responses to infant physical appearance. Infant Mental Health Journal. 1981; 2:56–61.

Hubel DH, Wiesel TN. Receptive fields of single neurones in the cat's striate cortex. Journal of Physiology. 1959; 148:574–591. [PubMed: 14403679]

Hugenberg K. Social categorization and the perception of facial affect: Target race moderates the response latency advantage for happy faces. Emotion. 2005; 5:267–276. [PubMed: 16187863]

Isbell LA. Snakes as agents of evolutionary change in primate brain. Journal of Human Evolution. 2006; 51(1):1–35. DOI: 10.1016/j.jhevol.2005.12.012 [PubMed: 16545427]

Jayakumar J, Sujata R, Dreher B, Martin PR, Vidyasagar TR. Multiple pathways carry signals from short-wavelength-sensitive ('blue') cones to the middle temporal area of the macaque. J Physiol. 2013; 591(1):339–352. 2013. [PubMed: 23070701]

Jones BC, DeBruine LM, Little AC, Conway CA, Feinberg DR. Integrating gaze direction and expression in preferences for attractive faces. Psychological Science. 2006; 17:588–591. [PubMed: 16866744]

Jones BC, DeBruine LM, Main JC, Little AC, Welling LL, Feinberg DR, Tiddeman BP. Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. Proceedings of the Royal Society B: Biological Sciences. 2010; 277:617–624. [PubMed: 19864283]

Kanwisher N. Functional specificity in the human brain: A window into the functional architecture of the mind. Proceedings of the National Academy of Sciences. 2010; 107:11163–11170.

Kanwisher NG, McDermott J, Chun MM. The fusiform face area: A module in human extrastriate cortex specialized for face perception. Journal of Neuroscience. 1997; 17:4302–4311. [PubMed: 9151747]

Kaas J. TITLE. In Cognitive Neuroscience, A reader. Ed Gazzaniga. 2000

Keltner, D., Ekman, P., Gonzaga, GC., Beer, J. Facial expression of emotion. In: Davidson, RJ.Scherer, KR., Goldsmith, HH., editors. Handbook of affective sciences. New York: Oxford University Press; 2003. p. 415-431.

Kveraga K, Boshyan J, Bar M. Magnocellular projections as the trigger of top-down facilitation in recognition. Journal of Neuroscience. 2007; 27:13232–13240. [PubMed: 18045917]

Kveraga K, Ghuman AS, Bar M. Top-down predictions in the cognitive brain. Brain and Cognition. 2007 Nov; 65(2):145–68. [PubMed: 17923222]

Kveraga, K., Boshyan, J., Bar, M. The proactive brain: Using memory-based predictions in visual recognition. In: Dickinson, S.Tarr, M.Leonardis, A., Schiele, B., editors. Object Categorization: Computer and Human Vision Perspectives. Cambridge University Press; 2009.

Kveraga K, Ghuman AS, Kassam KS, Aminoff E, Hamalainen MS, Chaumon M, et al. Early onset of neural synchronization in the contextual associations network. Proceedings of the National Academy of Sciences. 2011; 108(8):3389–3394.

LeDoux, JE. The emotional brain. New York, NY: Simon & Schuster; 1996.

Le Gal PM, Bruce V. Evaluating the independence of sex and expression in judgments of faces. Perception and Psychophysics. 2002; 2:230–243.

Maclin OH, Malpass RS. Racial categorization of faces: The ambiguous race face effect. Psychology, Public Policy, and Law. 2001; 7:98–118.

Maior RS, et al. Superior colliculus lesions impair threat responsiveness in infant capuchin monkeys. Neurosci Lett. 2011; 504(3):257–260. [PubMed: 21970966]

Marsh AA, Adams RB Jr, Kleck RE. Why do fear and anger look the way they do? Form and social function in facial expressions. Personality and Social Psychology Bulletin. 2005; 31:73. [PubMed: 15574663]

McKone E, Brewer JL, MacPherson S, Rhodes G, Hayward WG. Familiar other-race faces show normal holistic processing and are robust to perceptual stress. Perception. 2007; 36:224. [PubMed: 17402665]

Meeren HJM, van Heijnsbergen CCRJ, de Gelder B. Rapid perceptual integration of facial expression and emotional body language. Proceedings of the National Academy of Sciences. 2005; 102:16518–16523.

Michel C, Rossion B, Han J, Chung C, Caldara R. Holistic processing is finely tuned for faces of one's own race. Psychological Science. 2006; 17:608–615. [PubMed: 16866747]

Mogg K, Garner M, Bradley BP. Anxiety and orienting of gaze to angry and fearful faces. Biological Psychology. 2007; 76:163–169. [PubMed: 17764810]

Morand S, Thut G, Grave de Peralta R, Clarke S, Khateb A, Landis T, Michel CM. Electrophysiological evidence for fast visual processing through the human koniocellular pathway when stimuli move. Cereb Cortex. 2000; 10:817–825. [PubMed: 10920053]

Morris JS, Öhman A, Dolan RJ. Conscious and unconscious emotional learning in the human amygdala. Nature. 1998; 393:467–70. [PubMed: 9624001]

Nakayama, N. Vision going social. Adams, RB., JrAmbady, N.Nakayama, K., Shimojo, S., editors. Oxford University Press; p. xv-xxix.The science of social vision

O'Toole AJ, et al. The perception of face gender: the role of stimulus structure in recognition and classification. Memory & Cognition. 1998; 26:146–160. [PubMed: 9519705]

Pauker K, Ambady N, Weisbuch M, Sommers SR, Adams RB Jr, Ivcevic Z. Not so black and white: Memory for ambiguous group members. Journal of Personality and Social Psychology. 2009; 96:795–810. [PubMed: 19309203]

Pessoa L, Adolphs R. Emotion processing and the amygdala: from a "low road" to "many roads" of evaluating biological significance. Nature Reviews Neuroscience. 2010; 11:773–83. [PubMed: 20959860]

Peters J, Daum I, Gizewski E, Forsting M, Suchan B. Associations evoked during memory encoding recruit the context-network. Hippocampus. 2009; 19:141–151. [PubMed: 18777560]

Pineda JA, Sebestyen G, Nava C. Face recognition as a function of social attention in non-human primates: an ERP study. Brain Res Cogn Brain Res. 1994; 2:1–12. [PubMed: 7812173]

Redican, WK. An evolutionary perspective on human facial displays. In: Ekman, P., editor. Emotion in the Human Face. 2nd. Cambridge Univ. Press; Cambridge: 1982. p. 212-280.

Richeson JA, Todd AR, Trawalter S, Baird AA. Eye-gaze direction modulates race-related amygdala activity. Group Processes and Intergroup Relations. 2008; 11:233–246.

Rule NO, Ambady N, Adams RB Jr, Macrae CN. Us and them: Memory advantages in perceptually ambiguous groups. Psychonomic Bulletin & Review. 2007; 14:687–692. [PubMed: 17972734]

Sapir E. Speech as a personality trait. American Journal of Sociology. 1927; 32:892–905.

Schaefer ES, Plutchik R. Interrelationships of emotions, traits, and diagnostic constructs. Psychological Reports. 1966; 18:399–410.

Shepherd SV, Deaner RO, Platt ML. Social status gates social attention in monkeys. Current Biology. 2006; 16:R119–120. [PubMed: 16488858]

Shriver E, Young S, Hugenberg K, Bernstein M, Lanter J. Class, race, and the face: Social context modulates the Cross-Race Effect in face recognition. Personality and Social Psychology Bulletin. 2008; 34:260–278. [PubMed: 18212334]

Siegel S. Cognitive Penetrability and Perceptual Justification. Noûs. 2011; 46:201–222.

Tamietto, M., Weiskrantz, L., Geminiani, G., de Gelder, B. The Medium and the Message: Non-conscious processing of emotions from facial expressions and body language in blindsight; Paper presented at the Cognitive Neuroscience Society Annual Meeting; New York, NY: 2007.

Tanaka JW, Kiefer M, Bukach CM. A Holistic Account of the Own-Race Effect in Face Recognition: Evidence from a Cross-Cultural Study. Cognition. 2004; 93:B1–B9. [PubMed: 15110726]

Townsend JT, Wenger MJ. A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. Psychological Review. 2004; 111:1003–1035. [PubMed: 15482071]

Trawalter S, Todd A, Baird AA, Richeson JA. Attending to threat: Race-based patterns of selective attention. Journal of Experimental Social Psychology. 2008; 44:1322–1327. [PubMed: 19727428]

Tudusciuc O, Nieder A. Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans. Experimental Brain Research. 2010; 207:221–31. [PubMed: 20972775]

Turk DJ, Handy TC, Gazzaniga MS. Can perceptual expertise account for the own-race bias in face recognition? A split-brain study. Cognitive Neuropsychology. 2005; 22:877–883. [PubMed: 21038280]

Vuilleumier P, Armony JL, Driver J, Dolan RJ. Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nature Neuroscience. 2003; 6:624–631. [PubMed: 12740580]

Weisbuch M, Adams RB Jr. The functional forecast model of emotional expression processing. Social and Personality Psychology Compass. 2012; 6:499–514.

Weisbuch M, Pauker K, Lamer S, Adams RB Jr. Social Stratification and Reflexive Gaze-Following. under review.

Wiesenfeld AR, Klorman R. The mother's psychophysiological reactions to contrasting affective expressions by her own and an unfamiliar infant. Developmental Psychology. 1978; 14(3):294–304.

Wiggins JS, Broughton R. A geometric taxonomy of personality scales. European Journal of Personality. 1991; 5:343–365.

Zebrowitz. Reading Faces: Window to the Soul?. Boulder: Westview Press; 1997.

Zebrowitz L, Kikuchi M, Fellous JM. Facial resemblance to emotions: group differences, impression effects, and race stereotypes. Journal of Personality and Social Psychology. 2010; 98:175–189. [PubMed: 20085393]
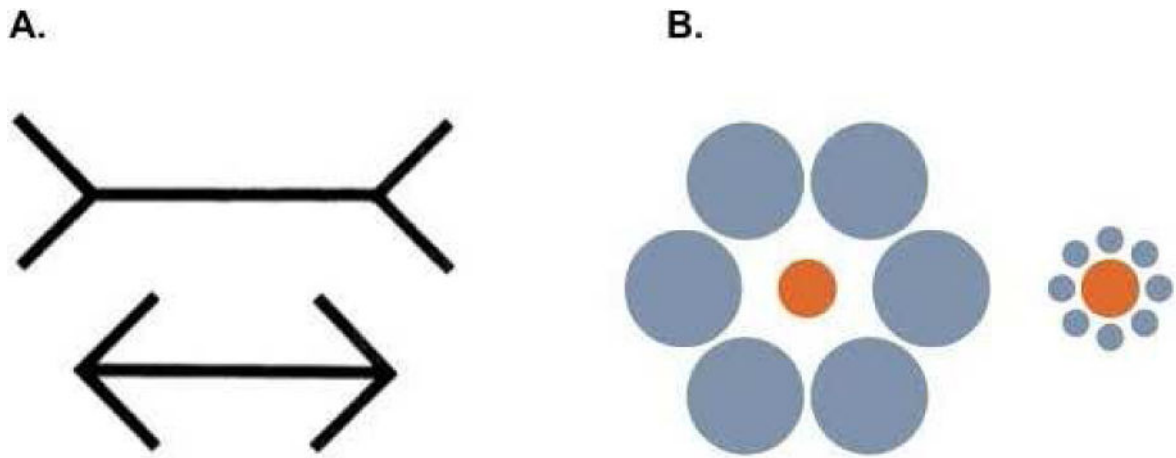
A.

B.

Figure 1.
Visual illusions often cited as examples of domain-specific, cognitively impenetrable visual processing: a) Müller-Lyer illusion and b) Ebbinghaus illusion.