The fields of social perception and emotion perception are motivated by an impressive everyday phenomenon: human perceivers are able to effortlessly, seemingly automatically, obtain a wealth of important information from visual perceptions of another person. From social categories to emotion to personality traits, these attributes are consequential for rapidly determining how to behave toward and around a given individual, so it is not surprising that the perceptual system is exquisitely sensitive to affectively and socially salient stimuli.

A great deal of research has outlined the brain regions that are specifically responsive to socially relevant stimuli like faces and bodies, including regions' sensitivity to social characteristics of these stimuli (e.g. social categories, emotions, and personality traits). While early visual processing regions in occipital cortex are responsive to the individual visual features of faces, higher-order visual processing regions such as the fusiform gyrus (FG) are more responsive to configurations of facial features, corresponding to more abstract properties such as categorical distinctions [53,31,28]. Leading from these important foundations, researchers have found that the FG may be involved in representing social information content of faces, including social categories such as sex [33,23], race [14], and emotion [67]. A recent study also observed these effects while controlling for models of low-level features of the face stimuli used [56], strengthening the evidence that the FG represents faces' more abstract social information.

Emotion perception relies more on dynamic cues for successful recognition, and much work has suggested the superior temporal sulcus (STS) plays an important role in this process [55,28,70]. This region is more generally sensitive to dynamic facial cues [28] and biological motion (i.e. walking and other naturalistic movements; [27]), and may serve to integrate multimodal information in social perception [51,50]. Despite the involvement of these specific regions, multi-voxel pattern analysis (MVPA) suggests that considerable information about faces or bodies is represented in a distributed fashion widely across ventral-temporal regions [28]. By and large, such work has taken a bottom-up approach, characterizing the perceptual mechanisms underlying the extraction of information from faces or bodies.

Beyond perceptions of social categories and emotion expressions, certain neural regions are also highly sensitive to the personality trait information conveyed by faces. Behavioral studies have long demonstrated that traits such as trustworthiness are reliably and consistently perceived from specific sets of facial features [48]. Neuroimaging evidence shows the amygdala responds to facial trustworthiness, with some amygdala subregions showing linearly increasing activation for faces appearing less trustworthy [70,19], consistent with the amygdala's role in vigilance for threats [35]. Other subregions show increasing activation for faces appearing more trustworthy or untrustworthy relative to neutral [54,61], consistent with the amygdala's role in processing affective significance in general (regardless of valence) and responding to motivationally relevant stimuli [15]. Both types of amygdala responses are observed even when faces are presented subliminally, demonstrating the amygdala's ability to draw even high-level social information from faces rapidly and automatically, even in the absence of awareness [24]. In terms of the visual processing of facial features associated with categorizations of trustworthiness (rather than rapid processing of their salience/affective meaning), the

**ABSTRACT**

The visual system is able to extract an enormous amount of socially relevant information from the face, including social categories, personality traits, and emotion. While facial features may be directly tied to certain perceptions, emerging research suggests that top-down social cognitive factors (e.g., stereotypes, social-conceptual knowledge, prejudice) considerably influence and shape the perceptual process. The rapid integration of higher-order social cognitive processes into visual perception can give rise to systematic biases in face perception and may potentially act as a mediating factor for intergroup behavioral and evaluative biases. Drawing on neuroimaging evidence, we review the ways that top-down social cognitive factors shape visual perception of facial features. This emerging work in social and affective neuroscience builds upon work on predictive coding and perceptual priors in cognitive neuroscience and visual cognition, suggesting domain-general mechanisms that underlie a social-visual interface through which social cognition affects visual perception.
FG likely plays an important role [59]. The findings in Freeman et al. [24] occurred in the absence of corresponding FG activity, making it unclear whether the subliminal amygdala effects were due to rapid responsivity in response to early processing in the FG, or potentially a subcortical route sensitive to specific features (e.g. [25,68]). As noted in the study, methodological factors such as using face rather than pattern masks may have also obscured FG effects, as this could induce adaptation effects and the FG in particular is sensitive to visual adaptation of face stimuli. The amygdala exhibits varied contributions to face processing, responding to facial attractiveness [36,69] as well as general aspects of facial structure like distinctiveness [39] and typicality [54]. Indeed, evidence shows that the amygdala is responsive to faces in general regardless of expression [40]. Future work is necessary to isolate the specific roles of amygdala subregions to explain these diverse findings.

Although the vast majority of work in social and cognitive neuroscience has focused on the determinant cues and bottom-up mechanisms that underlie extraction of social information, recent work suggests that this process is sometimes considerably influenced by higher-order social cognitive factors harbored in the perceiver, including prior knowledge (e.g., stereotypes or conceptual knowledge), attitudes, and motives [22]. According to current computational models such as the Dynamic Interaction (DI) Model, upon viewing another person’s face, social and emotion categories become automatically activated, which in turn automatically activate related conceptual associations [20]. Such conceptual associations then effectively become implicit predictions that shape the course of perception, biasing it to fall in line with those predictions [21,22]. Similar effects emerge not only for conceptual associations linked to a category, but also linked to a specific exemplar. For example, biographical and behavioral information about a person (particularly with affective or moral content) can influence attention and perception [1] and shape representations in the FG (e.g. [65]). The recurring conclusion in this line of work is that higher-order social cognitive processes may be seamlessly integrated into perceptual processing, fundamentally shaping social perception. Moreover, these effects are not irregular biases but fundamental characteristics of social perception.

The notion that higher-order cognitive processes may exert a top-down impact on perception has recently become important in the emotion literature as well. Although traditional work often assumed a direct “read-out” of universally recognized facial expressions [17], recent research reveals emotion perception to be particularly sensitive to external factors in the environment and harbored in the perceiver. For example, contextual information from the body and voice appear to dominate input from the face during emotion perception, such that when facial cues are incongruent with vocal and bodily expressions, the ultimate categorizations are congruent with the contextual input rather than the face itself [64,16]. Recent theoretical insights [9,37] and behavioral research [26,45] demonstrate the importance of emotion concept knowledge to emotion perception, suggesting that facial emotion expressions are relatively ambiguous until they are implicitly conceptualized as an instance of a specific emotion category.

Thus, the aforementioned work suggests that the visual perception of social and emotion categories may be flexibly shaped by higher-order cognitive processes. At the neural level, research in cognitive neuroscience has provided clues to the specific mechanisms by which such top-down impacts may occur. For instance, the object recognition literature demonstrates a role for the orbitofrontal cortex (OFC) in visual perception such that it is recruited when incoming visual input matches a pre-existing representation in memory or a task-based prediction, allowing it to potentially take on some of the visual processing load by supplying lower-level regions with visual expectations about category membership [58]. In particular, the medial OFC (mOFC) seems responsible for making connections between visual input and associations in memory [5,18]. Coarse, low spatial frequency input is sufficient to drive this expectation and prediction-based activity in the OFC, suggesting that the OFC exerts these visual predictions about category membership before categorization is complete [8].

Moreover, in the context of social perception of faces, multi-voxel patterns within the mOFC exhibit a representational structure of social categories that is shaped by stereotype knowledge (a representational structure that is also shared by the FG, potentially reflecting top-down modulation by the OFC; [56]). Representations in the mOFC are also more reflective of subjective categorical distinctions (male/female) rather than objective physical differences in a face’s gendered features [23]. Patients with mOFC/ventromedial prefrontal cortex lesions also have impaired access to gender stereotypes [42] and the mOFC has additionally been implicated in the implicit retrieval of stereotypes and person-based knowledge more generally [43]. Altogether, such work suggests that the OFC supplies automatic visual predictions, including those based on social concepts and stereotypes, to ventral-temporal regions involved in face perception. For example, early processing of cues to a social category like “female” or “Asian” may activate visual predictions that are based on stereotypes, facilitating the perception of stereotype-congruent visual information during the perceptual process. Our perspective is that this need not reflect some special process about stereotyping which the OFC is involved in, but rather reflects a domain-general characteristic of an interactive perceptual system that relies on top-down predictions or perceptual ‘priors’ from the OFC [22]. Indeed, much of this work dovetails with findings on predictive coding and top-down factors such as attention, context, and expectation in other perceptual processes (e.g. [7,34,58]). While this is often viewed as adaptive in the context of object recognition (e.g., facilitating recognition of an object in a congruent context that makes the object’s presence more likely), in the context of social perception the context can include stereotypical associations and the top-down signals on perception can in many cases be considered problematic and mala-daptive.

If the OFC is able to utilize expectations based on stereotypes or other conceptual associations to form implicit visual predictions that modulate ventral-visual representation, it is likely that the anterior temporal lobe (ATL) may be an important source of those associations. The ATL is consistently implicated in the storage and retrieval of semantic information and is considered a “hub” in semantic processing [49,10]. However, some researchers have argued that the ATL has a specialized role in storing socially relevant semantic knowledge [46]. For example, the ATL responds to faces when those faces are associated with prior knowledge [71], and some evidence suggests that the role of the ATL is to store and retrieve amodal representations of identity [2] and social knowledge more generally (such as the names, traits, and biographical details of known others; [46,13]). Together, these findings suggest a network wherein stereotypes and social knowledge from the ATL are rapidly retrieved by the OFC and integrated into ongoing processing of facial cues in the FG, enabling higher-order social cognitive factors to impact perceptual representation of visual stimuli.

Recent work suggests that such a social-visual interface may exist, particularly in the context of stereotypes. Stereotypes are generalized conceptual knowledge about social categories, which need not be personally endorsed and can be implicitly accessed. While researchers have long assumed that social categories spontaneously activate related stereotypes, the novel idea stemming from the perspective of the DI Model is that such activated stereotypes can shape earlier category activation itself, thereby shifting perceptions. Recent neuroimaging work has demonstrated how pairs of categories from different social dimensions (such as race and sex, or race and emotion) can become biased toward one another at the level of perceptual representation when those categories share stereotypical associations. In particular, researchers found that neural response patterns in response to sex (Male, Female), race (Black, White, Asian) and emotion (Happy, Angry) categories were systematically more similar at the level of visual perception when those categories were more similar at the level of stereotype knowledge, even after controlling for the visual similarity of the
face stimuli in question [56]. Specifically, the extent to which a participant stereotypically held categories such as ‘Black’ and ‘Angry’ to be similar (e.g., both associated with hostility) predicted a corresponding perceptual bias (e.g., non-Angry Black faces were temporarily perceived as Angry, measured with a mouse-tracking technique). In turn, such biases were reflected in the multi-voxel representational structure of FG and OFC patterns, such that neural response patterns for the ‘Angry’ and ‘Black’ categories in these regions exhibited greater similarity, controlling for possible bottom-up visual similarity in the faces themselves. As discussed earlier, the FG is centrally involved in high-level face perception and the OFC is involved in potentially providing visual predictions that can modulate such FG face representation [22]. The fact that stereotypes were able to shape the structure of perceptual representations of faces’ social categories in the FG suggests that higher-order social cognitive processes can bias relatively low-level perceptual representations.

Once these temporary biases are activated, there is some possibility for implicit control to occur as well. Additional neuroimaging work has investigated the impact of stereotypes on emotion perception in a racial context. Behavioral work consistently shows that Black targets are biased to be perceived as Angry, in line with the stereotype in the U.S. that Black individuals are aggressive and hostile [30]. In one study, researchers found that the dorsal anterior cingulate cortex (dACC), a region implicated in conflict monitoring as well as ambiguity processing [11,44] was responsive to the stereotypic incongruency of targets (e.g. happy Black and angry White faces). The DACC also showed greater functional connectivity (temporal correlation) with the FG during stereotypically incongruent trials, potentially suggesting top-down modulation of perceptual representations to help resolve the incongruency [29]. While the DACC and contiguous pre-supplementary motor area are involved in a wide range of functions, more recent work used a similar paradigm and showed that the DACC tracked similar social category conflicts when perceiving faces on a trial-by-trial basis as measured with mouse-tracking (which indexes response competition), and DACC activity again showed a functional relationship with the FG suggesting top-down assistance in resolving the perceptual conflict [57].

In addition, Heathman et al. [29] showed that subjects who were high in stereotypic associations between race and emotion exhibited greater activity in the dorsolateral prefrontal cortex (dIPFC) during stereotypically incongruent trials, such as a happy Black face. The dIPFC is implicated in cognitive control and response inhibition [38], among several other functions, suggesting inhibition of an inaccurate stereotype-driven association to allow veridical perception of a face (e.g. inhibiting stereotype-driven ‘Angry’ perception for a happy Black face). These results are particularly interesting as they emerged spontaneously during a passive viewing task, with no explicit goal for participants to attend to race or emotion. The pattern of results suggests the possibility that conflict-monitoring and control processes may come into play to help correct temporary top-down biases in social perception once they are set into motion. But it is important to note that the DACC and dIPFC are implicated in many related processes beyond conflict-monitoring and control, respectively, and more research is needed to clarify their involvement in the initial moments of social perception.

Research also suggests that an individual’s prejudice and evaluative biases toward particular social groups can impact visual perception of them. For example, MVP reveals that implicitly prejudiced individuals show more distinct (i.e. dissimilar) patterns of activity for own-race vs. other-race faces when compared with individuals who do not show such a bias [12]. Additionally, the FG shows a distinction between in-group and out-group members (with greater activity for in-group rather than out-group members, potentially reflecting enhanced processing and encoding), even when the groups in question are “minimal” groups invented in the experimental setting, suggesting that group-related top-down effects in the FG reflect higher-level encoding of faces beyond familiarity with a particular group’s visual features [63]. In some cases, it is possible that these in-group/out-group distinctions may supersede other more visually salient categories such as race, but existing results are mixed. For example, participants show implicit favoritism [62] and enhanced neural processing in the FG [62] for minimal in-group members even when these groups contain equal numbers of racial in-group and out-group members. However, multivariate approaches show that race can still be classified by different neural patterns for own- vs. other-race faces (within the same minimal group) in the FG [52], and that categorical distinctions of group and race in the FG are sensitive to current processing goals (e.g. task-based shifts in attention; [32]). Together, these results suggest that individual differences in prejudicial biases and intergroup effects are present at the level of visual representation and may manifest as perceptual biases.

Conclusion

Together, the research reviewed here suggests that social and emotion perceptions are not direct products of facial features but rather seamlessly influenced by higher-order social cognitive processes, including stereotypes, social-conceptual knowledge, and evaluative biases. Existing work in cognitive neuroscience has extensively documented the way that conceptual knowledge and top-down expectations can be rapidly integrated into perceptual experience. Work in social and affective neuroscience is beginning to converge with and extend these findings, ultimately suggesting a domain-general system in which implicitly activated top-down expectations are used to help make sense of incoming visual input. This may be carried out by a network permitting such flexible impacts on social perception, whereby the OFC generates visual predictions informed by associations retrieved from the ATL to bias FG representations of faces in line with those predictions [22]. However, as this work is only in its infancy, a number of questions remain. Functional and structural connectivity approaches will be important to better characterize the functional relationships among regions in this putative network, including more recent multivariate functional connectivity approaches [3,4]. Existing work with connectivity approaches describes rich anatomical connectivity between the FG, ATL, and OFC, scaffolding a flexible and dynamic neural system responsive to the high processing demands of person perception (e.g., [6,41,47,66]). Advances can also be made to more fully delineate the roles of automaticity and control in top-down impacts on perception, including whether conscious goals or regulatory strategies may be used to reduce top-down impacts of stereotypes that are undesirable for individuals motivated to avoid bias. Finally, due to space constraints, we are unable to offer a diligent review of research on whether facial features are accurate signals for trait attributions and other inferences such as sexual orientation (for a recent discussion, see [60]). However, the work we review here shows that regardless of whether facial features can permit accurate or inaccurate judgments with respect to a target’s dispositional qualities, the perceptual system brings a great deal to bear on their processing and interpretation. We look forward to continued research in this area to shed light on the social-visual interface and the complex interdependence between social cognition and visual perception.

Acknowledgement

This work was funded in part by NSF grants BCS-1423708 and BCS-1654731 to J.B.F.

References
