The Social-Emotional Processing Stream: Five Core Constructs and Their Translational Potential for Schizophrenia and Beyond

Kevin N. Ochsner

Background: Cognitive neuroscience approaches to translational research have made great strides toward understanding basic mechanisms of dysfunction and their relation to cognitive deficits, such as thought disorder in schizophrenia. The recent emergence of Social Cognitive and Affective Neuroscience has paved the way for similar progress to be made in explaining the mechanisms underlying the social and emotional dysfunctions (i.e., negative symptoms) of schizophrenia and that characterize virtually all DSM Axis I and II disorders more broadly.

Methods: This article aims to provide a roadmap for this work by distilling from the emerging literature on the neural bases of social and emotional abilities a set of key constructs that can be used to generate questions about the mechanisms of clinical dysfunction in general and schizophrenia in particular.

Results: To achieve these aims, the first part of this article sketches a framework of five constructs that comprise a social-emotional processing stream. The second part considers how future basic research might flesh out this framework and translational work might relate it to schizophrenia and other clinical populations.

Conclusions: Although the review suggests there is more basic research needed for each construct, two in particular—one involving the bottom-up recognition of social and emotional cues, the second involving the use of top-down processes to draw mental state inferences—are most ready for translational work.

Key Words: Amygdala, cingulate cortex, cognitive neuroscience, emotion, prefrontal cortex, schizophrenia, social cognition, translational research

From time to time, we all fail to respond adaptively to life’s challenges. For individuals with clinical disorders, however, these failures may be chronic and pervasive. An essential goal of behavioral and neuroscience research is to understand how and why this happens. One influential approach has been to use basic cognitive neuroscience models to describe how and when clinical symptoms arise from dysfunction in core mechanisms of attention, memory, and other higher cognitive processes. This translational approach has taken basic cognitive neuroscience models of prefrontal function and applied them to the study of positive symptoms in schizophrenia. This work has shown, for example, that individuals with schizophrenia show disorder-specific behavioral deficits in maintaining task contexts that both predict thought disorder symptoms and resolve with treatment (1–4). This work has been less successful, however, in explaining the social and emotional dysfunctions that characterize negative symptoms in schizophrenia and many DSM Axis I and II disorders more broadly (2).

The rapid development of social cognitive and affective neuroscience (SCAN) as distinct disciplines (5–7) offers opportunities for these kinds of translational bridges to be built. The proliferation of new SCAN findings is both a blessing and curse for basic and clinical neuroscientists, however. On one hand, new findings can provide material for building new kinds of bridges (8,9). On the other hand, with the multiple approaches and methods this new work employs, it can be difficult to figure out how diverse pieces of data fit together into core neurofunctional constructs. Identifying these constructs is essential, because our theoretical models of them determine what scientific questions we ask about their basic nature and translational potential. Given that performance on behavioral measures of social cognition and emotion might predict functional outcomes in schizophrenia (10–15), the time is ripe for neuroscience research to examine the brain systems underlying these abilities in schizophrenia and beyond.

The overarching goals of this article are to distill a set of key constructs from the growing data on the neural bases of social and emotional abilities that can be used to generate questions about the mechanisms underlying negative symptoms in schizophrenia and, by extension, clinical disorders of emotion more generally. Toward these ends this article has two parts. The first briefly sketches a framework in which five constructs comprise a social-emotional processing stream. The second considers how future basic research might flesh out this framework and translational work might relate it to the study of negative symptoms in schizophrenia and other clinical disorders. In this regard the article was motivated by the need to provide a framework for the CNTRICS (Cognitive Neuroscience for Treatment Research to Improve Cognition in Schizophrenia) initiative (2), which is concerned with adapting measures from cognitive, social, and affective neuroscience for use in clinical trials in schizophrenia.

The Social-Emotional Processing Stream

The basic premise of this framework is that, in many if not all cases, human social and emotional behaviors are highly intertwined. Consider, for example, how a social cognitive or an affective neuroscientist might study different aspects of a social interaction. The social cognitive neuroscientist might focus on how each person draws inferences about the momentary thoughts and feelings of their interaction partner as well as...
partner’s enduring traits and tendencies. The affective neuropsychologist might focus on each person’s emotional response, how each person regulates it, and how each partner identifies each other’s emotional expressions. Although these social and emotional questions have historically been the province of different disciplines, are the phenomena of interest completely distinct? This review argues that they are not: how you assess the intentions (e.g., aggressive) and dispositions (not usually that way) of another person is part of the appraisal process that assesses what emotion that person is expressing and determines your emotional response to the person (e.g. fear) as well as how you might regulate that response (e.g., judging the aggression to be circumstantial) (for discussion see 16).

The common intertwining of social cognitive and affective phenomena makes sense, given that many researchers believe emotions arise from appraisals of the goal relevance of a stimulus and that people are typically the most goal-relevant stimuli in our daily lives. This is not to say that we cannot experience emotions in non-social contexts (e.g., disgust at trash) but rather that it is difficult to have social interactions without emotion. This might explain why the paradigms used in social cognitive and affective neuroscience research are strikingly similar: ostensibly social cognitive tasks often involve affective processes (including attitudes), and ostensibly affective tasks often use social stimuli (like faces or social images). It might also explain why functional imaging and lesion studies of social cognitive and affective phenomena consistently implicate a common set of brain systems (16).

With this in mind, this article uses the term “social-emotional processing stream” to refer to the set of psychological and neural processes that encode socially and emotionally relevant inputs, represent their meaning, and guide responses to them. The sections that follow sketch five core constructs that are the key constituents of this stream. Selection of these constructs was guided by two factors. First, human and animal data had to suggest that there are reliable neural correlates of the ability/construct in question. Second, theoretical models of social cognition and emotion (16–21) were used to guide grouping of behavioral phenomena under each construct.

The end product is the heuristic model illustrated in Figure 1, where the term “construct” refers to categories of social cognitive and affective abilities that are valid and distinct in so far as they have been tied to distinct but related sets of neural systems. These constructs lie along a rough hierarchy of processes engaged when we initially learn the value of a stimulus (Construct 1); subsequently re-encounter it and recognize its value (Construct 2); understand the beliefs and feelings of a person stimulus—that could be oneself—in a bottom-up, experiential (Construct 3) or top-down, attributional manner (Construct 4); or try to regulate responses to a stimulus in a context appropriate manner (Construct 5). Here, “value” refers to whether a given stimulus is good or bad or should be approached or avoided, whereas “response” refers to the behaviors we measure as evidence that this value has been computed. Because current data do not allow us to clearly disentangle the neural correlates of the valuation and response stages, these two terms are often used in combination here.

Construct 1: Acquisition of Social-Affective Values and Responses

The first construct concerns the universal need to learn which stimuli and actions—whether social or non-social—lead to aversive as opposed to appetitive outcomes. For decades, acquisition of social-affective values and responses has been studied in simple animal models of conditioning and reward learning that only recently have been extended to humans with functional imaging and patient studies. Together, these data provide perhaps the strongest evidence for any of the proposed constructs.

The two neural systems most strongly implicated in affective learning—the amygdala and striatum—are evolutionarily old subcortical structures that receive multi-modal perceptual inputs.

![Diagram](Image)

**Figure 1.** Diagrammatic illustration of the relationships between five proposed core abilities/constructs for social and emotional behavior. Examples of each are listed underneath the box naming each ability/construct. For illustrative purposes a linear flow of information is shown between the systems underlying each proposed ability/construct, although in principle the systems underlying each construct might operate independently or in various combinations. See text for details.
and are interconnected with autonomic control centers and neuromodulatory systems (22–24) (Figure 2A). Classically “limbic” regions such as the medial prefrontal cortex (mPFC) and insula (see next section) also play key roles in affective learning via interconnection with the amygdala and striatum (Figures 2B and 2C).

The amygdala’s role in affective learning has been elaborated primarily with classical (aka Pavlovian) fear conditioning paradigms in which an initially neutral “conditioned” stimulus (CS; e.g., a tone) is repeatedly paired with an intrinsically aversive “unconditioned” stimulus (UCS; e.g., a shock). Over time, the CS comes to elicit behavioral “conditioned” responses (CRs; e.g., freezing) that might be similar to those initially elicited by the UCS. Elegant animal experiments have shown that the CS-UCS association involves interconnections between the basal and lateral amygdala nuclei and that the behavioral components of the CR depend on brainstem centers that receive projections from the basolateral complex via the central nucleus (25).

Human imaging and lesion studies have confirmed the role of the amygdala in classical fear conditioning (e.g., 26,27) and have extended animal work by showing that the amygdala is critical for acquiring conditioned fear responses to social stimuli that might act as CSs, such as faces or facial expressions of anger (28,29).

The ventral portions of the striatum (VS) are critical for learning which stimuli or behavioral responses predict rewarding or reinforcing outcomes (24). For example, with simple stimulus-reward association paradigms, single unit recording studies in nonhuman primates have shown that the function of the VS is well-described by a simple learning model in which dopamine release enables VS neurons to encode the timing of an expected reward, with release adjusted either upward or downward as a function of whether that expectation is met (24). Human imaging studies have corroborated this model by showing that VS activity increases when a participant anticipates or receives an unexpected monetary reward (30) and that it varies as a function of whether they have been led to expect a reward that occurs at an unexpected time (31,32).

Although the amygdala and VS play critical roles in learning which stimuli predict aversive and rewarding outcomes, respectively, these are not their only roles in acquiring affective responses. For example, the amygdala’s central nucleus might play a more general role in orienting attention to and encoding into memory affectively salient stimuli, which might relate to its role in signaling when the reward-related value of a stimulus changes (33) and consolidation of memories for affectively relevant experiences (34). Furthermore, interactions between the amygdala and VS might be critical for learning more complex affective associations (22,35,36).

Information the amygdala and VS send to the medial portions of the orbitofrontal and ventral mPFC (vmPFC) is important for representing the affective valence of stimuli as it is updated across contexts (36,37). In animal studies, orbitofrontal cortex (OFC) neurons fire in response to various kinds of motivationally relevant stimuli and update this firing more rapidly than the VS or amygdala as stimulus-reward associations change (37). In like fashion, human functional imaging studies have shown that vmPFC and OFC might respond to both rewarding and aversive outcomes and are sensitive to changing reward values (38–42).

Taken together, extant evidence suggests that the amygdala, VS, and vmPFC/OFC form a circuit essential for encoding the affective value of stimuli. One caveat to these data, however, is that a few studies have shown directly that these structures are important for acquiring the affective value of social stimuli, per se. Perhaps the most salient example is a recent study showing that the amygdala is essential for conditioned fear responses acquired by observing others undergo the conditioning procedure (43). Given these data, the connectivity of these systems, and data described in the next section that these structures respond to nonverbal social cues (such as faces) whose affective significance presumably has already been learned, it is safe to assume that simple affective learning systems are involved in social learning. As learning becomes more social, however—for example involving drawing inferences about mental states—additional structures such as dorsal regions of mPFC (see Construct 4) might also become important (43).

Construct 2: Recognition of and Response to Social-Affective Stimuli

Once the social-affective value of a stimulus has been learned, it is important that an organism can quickly identify it in the future and respond appropriately. The systems important for affective learning described in the preceding text and posterior cortical regions involved in representing nonverbal cues are important for this ability.

Perhaps the best-known finding in this domain is that the amygdala is critical for the recognition of stimuli that directly or indirectly signal the presence of a potential threat, such as the faces of seemingly untrustworthy individuals (44–46) and fearful facial expressions and the widened eyes and enlarged eye whites that uniquely characterize them (47,48). Topics of debate include the extent to which task factors, levels of anxiety and depression, and genetic factors determine the magnitude and attentional independence of the amygdala’s response to these fear cues (49–56).

Given that the amygdala also responds to novel and positive stimuli and that it is sensitive to the configural meaning of specific eye gaze/facial expression combinations (57), some have offered a broader conceptualization of the amygdala as a “surveillance” system that continuously monitors the environment for affectively relevant stimuli and modulates activity in perceptual and memory systems to detect and encode them (58,59). On this view, ambiguous and novel stimuli are (potentially) relevant until an organism learns otherwise (59).

The striatal and medial prefrontal systems described earlier also have been implicated in recognizing stimuli whose value they encode. Thus, imaging studies have shown that the VS and ventromedial PFC respond to the faces of attractive people (60,61) or consumer goods that one would like to purchase (62), presumably because of their learned (or perhaps innate) reward value. Ventromedial PFC and nearby regions of the anterior cingulate cortex also respond during like/dislike or preference judgments for various kinds of stimuli (63–65), presumably because indicating with a key press that one likes a stimulus is either an instance of expressing a preference one has already acquired or is an instance of learning that one has this preference.

Patterns of connectivity are also useful for understanding the functional roles in social-affective learning and recognition for each of these regions. For example, the fact that the amygdala receives multimodal perceptual inputs, including some that (at least in rodents) might provide quick inputs that bypass the cortex, suggests that this structure might be well suited for a role as a “surveillance” system. Another region whose pattern of connectivity might relate to its function is the insula, the cortical region connecting the temporal and frontal lobes that lies beneath the Sylvian Fissure (Figure 2C). The insula has been
described as viscerotopic map—with its posterior regions receiving ascending somatosensory information, including pain—that projects forward to anterior regions that are interconnected with frontal regions implicated in attention, control, and speech articulation (66–68). This mapping might explain why both regions might be activated by pain but only the anterior has been associated with the experience and recognition of facial expressions of disgust (69,70), an emotion that involves the oral expulsion of potential contaminants. Although some have argued that the anterior insula is critical for disgust (69), this has been questioned, because it also responds to other aversive facial expressions, memories, and images (71,72); might be active during classical conditioning (26); and is active (on the right) when one interoceptively detects one’s own heartbeat (73). These data motivate the view that the anterior insula plays a general role in negative affective experience (72,74).

Cortical regions around the superior temporal sulcus (STS) also play a role in the recognition of social/affective values. This was first identified in single unit recording studies in nonhuman primates and has since been extended in numerous functional imaging studies in humans. In human imaging studies, the STS responds to a variety of nonverbal cues that might include images of moving eyes, lips, mouths, grasping movements, and abstract stimuli that depict biologically plausible motion (75–77). The latter type of stimulus includes well-known point-light videos showing individuals walking, dancing, or engaging in other social or motivationally relevant actions (78). These regions lie just anterior to the temporal parietal junction (TPJ), which has been implicated in controlling the focus of attention (79) as well as in the representation of beliefs (80) (see Construct 4). The close proximity these regions might make sense, given that the perception of nonverbal cues, such as the direction of another person’s eye gaze, might automatically orient our attention in the direction that person is looking (81). Exactly how these regions communicate with one another, however, and the extent to which they represent perceptual, attentional, or higher-level semantic information is currently a topic of debate (82).

In sum, extant evidence strongly suggests that regions involved in learning the affective value of a stimulus also support recognition of it later on and that superior temporal regions are important for recognizing nonverbal social cues. Important questions remain, however, about how best to characterize the function of these regions. For example, some characterize the affective learning regions as having the specific function of recognizing a particular kind of stimulus (e.g., a fearful or disgust facial expression) (69), whereas others characterize them in terms of processes that are not domain-specific (e.g., 20,72).

Construct 3: Embodied Simulation or Low-Level Mental State Inference

There is more to understanding the meaning of a social-affective stimulus than simply being able to place it in the appropriate category as a fearful expression, a preferred product, or an attractive face. Indeed, theory and research suggests that, beyond such simple recognition judgments, the meaning or value of stimulus is embodied in our experience of it. In some sense, all of our experiences are embodied—we inhabit physical bodies that feel pain and whose palms sweat and muscles contract in the readiness for action. For present purposes, the key notion is that these responses are important components not just of our own direct first-person experience but might be used as “embodied simulations” that help us vicariously understand the experience of others as well (83–85).

Figure 3. The logic behind and regions implicated in studies of shared representations (see Construct 3). (A) Venn diagram showing the shared representation logic. The idea is that regions commonly activated during the first person experience of an action, pain, or emotion (left circle) might also be activated when observing others experiencing action, pain, or emotion (right circle). To the extent that these regions overlap, the information represented there might play a dual or shared role in supporting the direct experiential understanding of one’s own and others actions. (B) Semi-transparent lateral view of the right hemisphere showing inferior parietal and ventral premotor regions implicated in the putative “mirror system” for programming and recognizing intentional actions. Highlighted are anterior regions of the insula implicated in pain experience, pain empathy, disgust experience and disgust empathy, and negative affective experience more generally. See text for details. BA, Brodmann area.

Neuroscience data supporting this claim come from studies asking whether the neural systems involved in the execution of a motor act, the experience of pain or an emotion, also are active when a participant observes another person engaging in that same act or having the same kind of experience. As illustrated in Figure 3A, to the extent that common systems are involved, it has been argued that the perception of others is supported by or shares the same representations that support first-person experience (86,87).

The first data of this sort came from single unit recording studies in nonhuman primates showing that approximately 25% of neurons in the ventral premotor and inferior parietal cortex (Figure 3B) would fire when the animal performed an action as well as when it observed the experimenter or another animal performing an action with the same goal if not the identical means of execution (e.g., a different means of grasping a cup) (88). These “mirror neurons” were interesting because they seemed to encode the intention behind an action regardless of who performed it, and it was hypothesized that their activation could provide the basis for understanding the intentions behind the actions of another person. Human imaging research using the shared representation logic (Figure 3A) later provided for converging evidence for the existence of a similar human “mirror system” (89,90), although individual mirror neurons have yet to be observed directly.

Subsequent studies extended this logic to other domains where the activation of shared representations has been hypothesized to provide a basis for empathy. For example, numerous studies of pain empathy have shown the activation of two regions that receive ascending nociceptive inputs—the mid-
region of the anterior cingulate cortex (Figure 3B) and the anterior insula (Figure 3D)—when individuals directly experience and when they observe others experiencing physical pain (91–95). Similar findings were obtained in a study of disgust, which found that sniffing disgusting odors and watching others doing the same activated overlapping portions of the anterior insula (96) (Figures 2C and 3D).

The assumption in all of these studies is that common activity in the parietal, premotor, cingulate, and/or insular cortices provides the basis for the vicarious empathic experience and therefore understanding of another’s actions, pain, or disgust. This is consistent with recent findings that individuals with autism—who exhibit gross impairments of social behavior—might show reduced activation in the prefrontal portions of the so-called “mirror system” (97). There are a couple of problems with this assumption, however. First, the regions in question have been activated by a variety of motor actions and/or affective experiences, and at present it is not possible to determine whether common activity in the insula, for example, reflects the experience of disgust, pain, or some other kind of negative affective experience. Thus, it is possible that when I experience pain I am afraid, but when I see you experience pain I am disgusted. Second, to date, no studies have provided behavioral measures that could verify that direct and vicarious experience is similar or that activity in shared representation regions supports accurate judgments about and understanding of another person’s experience. For example, it would be desirable to show that activity in ventral premotor cortex or the mid-cingulate cortex might predict an individual’s ability to accurately judge the intentions behind an action or the nature of their painful experience. Although analytic and behavioral methods have been and are being developed that can address these issues, they have not yet been applied to imaging studies of shared representations. In future work, measures of behavioral mimicry or correlations between one’s self-reported experience and the experience that others judge you are having could address this issue (95,98–100).

Finally, it is important to note that the shared representation logic might be used to study not just how we understand others but how we recognize the meaning of their actions and respond to them as well. This has been shown in studies of the neural response to social rejection, which have shown that mid-cingulate and insula regions implicated in pain also are active when one experiences rejection (101). These data suggest that we understand what it means to be socially isolated in part by experiencing what it would be like to be physically hurt (102).

In sum, although extant data are consistent with the notion that the activation of shared representations enable us to simulate the experience of others, it is not yet clear when and how these simulations truly match the experience of others and enable us to accurately understand them. Another way of stating this is that the bottom-up, stimulus-driven activation of shared representations might support the vicarious understanding of another person’s experience, but the nature of that understanding remains to be determined (16). We do know that this understanding is low-level in the sense that the supporting systems represent the experiential properties of a stimulus rather than higher-level symbolic interpretations of it, which is considered in the next section.

Construct 4: High-Level Mental State/Trait Inference

One problem in interpreting the meaning of social stimuli is that they often are ambiguous. Take, for example, the image of a smiling face commonly used in many studies of facial expression recognition. The assumption is that the smile unambiguously communicates happiness. Anyone who has played cards or bought a used car knows that this is not the case, however, and that the meaning of a smile is determined by the context in which it is displayed. Importantly, it seems that in many cases the recognition and low-level motor and affective “simulation” processes described under Constructs 2 and 3 are insufficient for representing these complex types of intentional mental states (103,104). To understand them, we must use higher-level (possibly symbolic) representations of mental states to take into account situational/contextual information that constrains the meaning of a social action.

Perhaps the most well-studied example comes from studies of theory of mind (TOM) that employ variants of the false belief task. In this task participants read vignettes describing the actions of a character who possesses a false belief about the state of the world (105,106). The participant’s task is to correctly assess that belief. Because this judgment cannot be made on the basis of perceptual information, general knowledge about the physical world, or information that the participant herself knows to be the current true state of affairs, it is often considered to be the best test of individual’s ability to represent the mental states of other people (107). This task was originally developed to assess the developing child’s capacity to understand and explain the behavior of others in terms of internal mental states, such as their beliefs, desires, feelings, and goals. As the first task adapted to studying mental state inference in human functional imaging research, numerous studies have since employed vignette or even cartoon variants of it (105). In general, they have shown activation of a network of regions including dorsal and rostral mPFC and adjacent paracingulate cortex, the posterior cingulate/precuneus, temporal-parietal junction, STS, and the temporal pole. Sometimes referred to as the “mentalizing” network (108) (Figure 4), portions of this network—but most commonly the mPFC—also have been activated during other tasks that presumably rely upon the ability to infer mental states. These include playing strategic games against a human opponent (109–111), watching video clips of abstract shapes whose movements seem intentional (104,112,113), and forming or retrieving an impression of a person from a photograph of their face (114,115). The mPFC is the primary focus here because it is the most reliably activated across studies and the bulk of attention has been focused on unpacking its functional organization.

Intriguingly, some of the same mPFC regions implicated in mental state inference have also been implicated in accessing and making judgments about one’s own mental states and enduring traits. For example, judging your emotional response to

www.sobp.org/journal
a photograph activates regions of the mPFC also activated when judging the emotion of the people in that photo (116,117). Similarly, judging whether trait words describe you or close other might also activate common and mPFC regions (118). These data, like the data on embodied simulation and low-level mental state inference, suggest that some of the same processes used to make judgments about the self are used to make judgments about others. In this case, the processes are higher-level and involve the representation of belief states, which makes sense given that the mPFC is an integrative region that receives inputs from dorsal lateral and parietal regions implicated in working memory and spatial attention, as well as orbitofrontal regions that represent the motivational value of a stimulus (119). Intriguingly, mPFC sends projections to autonomic and endocrine centers that might enable current beliefs to influence visceromotor response channels, such as heart rate and galvanic skin response (120,121).

Self and other judgments might not depend upon entirely overlapping regions of the mPFC, however. Some experiments have shown, for example, that there might be distinct mPFC subregions associated with accessing information about self or other (114,117,118,122,123). Exactly how mPFC is organized with respect to making attributions about self or other is a current topic of debate. Some data suggest that ventral/perigenual as opposed to dorsal and rostral regions are more strongly associated with judgments about self and others, respectively (114,122). Other data suggest, however, that there might be other dimensions of organization related to self and other judgment that might explain this apparent difference (118,124). Consider, for example, that retrieving exemplars of affective categories (e.g., generating “machete” from the cue “weapon”) activates dorsal and rostral and mPFC (125,126), whereas expressing a preference for a stimulus or receiving a reward or punishment tends to activate more ventral regions of the mPFC (117). These data suggest that dorsal/rostral mPFC might be important for the explicit categorization of mental states (whether they are your own or someone else’s), whereas ventral mPFC provides a coarser representation of the motivational value of a stimulus that can guide action in the absence of explicit mental state attributions (16,117).

In sum, the mPFC clearly plays a key role in mental state inference, although the specific contributions of individual subregions to this ability remains to be clarified. That being said, the mPFC is by no means the only important component of a putative “mentalizing network.” Indeed, much recent attention has been focused on the roles of superior temporal regions in representing nonverbal visual cues that might provide clues to the intentions of others (see Construct 2) (75), of the TPJ in representing beliefs (106), of the precuneus in self awareness (127), and of the temporal pole in representing emotion knowledge (128). Unpacking the individual contributions to mental state inference of each of these regions—and whether and how similar systems are used for understanding one’s own mental states as opposed to the mental states of others—will be an important focus for future basic research (16,82).

Finally, it is worth noting that research on Constructs 3 and 4 are in many ways interrelated. Basic research on both constructs has been concerned with how we understand our own actions and experiences as well as those of others and the nature of the relationship between them. They differ, however, in the kinds of representations under investigation. Work on low-level mental state inference focuses on perceptual, motor, visceral, and affective representations that might support direct experiential understanding, whereas work on high-level mental state inference focuses on more abstract, semantic, and categorical representations that might support a symbolic or descriptive understanding of experience and action.

Construct 5: Context-Sensitive Regulation

The final construct concerns the ability to regulate one’s judgments about and behavior toward others in a context appropriate manner. As illustrated in Figure 5, this regulatory ability manifests itself in at least three ways, with each form of regulation differing in complexity and depending upon related but distinct sets of underlying neural systems (21).

The first can be termed “description-based regulation” because it involves the use of mental state inference, language, memory, and selective attention to reinterpret or reappraise the meaning of a social-affective stimulus (129,130). For example, one might explicitly reappraise an initially insulting remark if one could determine that it was in fact unintentionally hurtful. Here, one might use working memory to hold in mind a linguistic narrative about the other person’s mental states—while at the same time directing attention to their facial expressions and body movements to verify that the remark was meant to be playful and withholding the pre-potent tendencies to interpret their action as aggressive and respond in kind. Reappraisal has been studied by asking participants to reinterpret the meaning of affectively arousing photographs or anxiety-provoking situations in ways that either diminish or enhance their affective response (131). By and large, this work has shown that reappraisal depends upon activity in dorsal and lateral prefrontal regions implicated in language, attention, memory, and response selection (often collectively referred to as cognitive control) as well as in mPFC regions implicated in mental state inference (for reviews see 21,132). Activity in these control systems modulates activity in regions implicated in emotional responding, such as the amygdala or insula. Within these general constraints, the specific frontal regions activated across studies have varied considerably, however, which might have to do with the variability in the specific reappraisal strategies employed in each experiment (21,132). In addition, studies to date have focused primarily on negative affect, and less attention has been paid to the question of whether these same neural systems are used for regulating positive emotions or any single specific emotion.
A second regulatory ability might be termed “outcome-based regulation” because it depends on the re-mapping or re-learning of contingencies between stimuli or actions and affective outcomes. In contrast to description-based regulation, which depends upon high-level mental descriptions of the affective value of a stimulus, this form of regulation depends upon updating the value of the stimulus via direct experience with the affective outcomes associated with it. Perhaps the most well-studied example of this form of learning is extinction of the conditioned fear response. As described earlier, fear conditioning involves learning that an initially neutral stimulus (the CS) predicts the occurrence of an intrinsically unpleasant outcome (the UCS). During extinction the CS is repeatedly presented without the UCS. Over time, conditioned responses to the CS diminish as the organism learns that one no longer needs to fear that the unpleasant UCS will soon follow. Recording and lesion studies in animals as well as functional imaging studies in humans have implicated a region of the ventromedial/medial orbital frontal cortex in this ability (133,134).

On the basis of these results, some have characterized this mPFC region as having an inhibitory function. In the context of the work reviewed in the preceding text, however, it can be seen that this region is similar to those implicated in studies of reward learning, preference judgments, and certain kinds of social or self-reflective inference. Seen in this light, extinction learning can be seen as a form of updating or recontextualizing the affective value of the stimulus. This interpretation is consistent with other work implicating orbital regions in another variant of outcome-based regulation known as stimulus-reward reversal learning. In reversal learning experiments an individual is led to expect a rewarding outcome whenever one of two stimuli (e.g., A but not B) is selected. After this association is learned to criterion, the stimulus-reward association is reversed and stimulus B is now associated with the reward, whereas stimulus A is not. In animals and humans, lesions of ventromedial/medial OFC impair this ability (135–137), and these regions are active during imaging studies of reversal learning (138). In social contexts OFC lesions might manifest this deficit in properly evaluating the contextual value of a stimulus in interesting ways. Problems might include comments and actions that are inappropriately intimate or sexual, failing to appreciate social faux pas, and affect that is greater or lesser than the expected norm for a situation, especially when the unpleasant UCS will soon follow. Recording and lesion studies in animals as well as functional imaging studies in humans have implicated a region of the ventromedial/medial orbital frontal cortex in this ability (133,134).

On the basis of these results, some have characterized this mPFC region as having an inhibitory function. In the context of the work reviewed in the preceding text, however, it can be seen that this region is similar to those implicated in studies of reward learning, preference judgments, and certain kinds of social or self-reflective inference. Seen in this light, extinction learning can be seen as a form of updating or recontextualizing the affective value of the stimulus. This interpretation is consistent with other work implicating orbital regions in another variant of outcome-based regulation known as stimulus-reward reversal learning. In reversal learning experiments an individual is led to expect a rewarding outcome whenever one of two stimuli (e.g., A but not B) is selected. After this association is learned to criterion, the stimulus-reward association is reversed and stimulus B is now associated with the reward, whereas stimulus A is not. In animals and humans, lesions of ventromedial/medial OFC impair this ability (135–137), and these regions are active during imaging studies of reversal learning (138). In social contexts OFC lesions might manifest this deficit in properly evaluating the contextual value of a stimulus in interesting ways. Problems might include comments and actions that are inappropriately intimate or sexual, failing to appreciate social faux pas, and affect that is greater or lesser than the expected norm for a situation, especially when the unpleasant UCS will soon follow. Recording and lesion studies in animals as well as functional imaging studies in humans have implicated a region of the ventromedial/medial orbital frontal cortex in this ability (133,134).

The third regulatory ability, termed “choice-based regulation,” involves weighing the relative values of choice options to balance short-term versus long-term gains. In this form of regulation, the act of making a choice to favor one type of gain or the other has a de-facto regulatory effect upon behavior. The classic example comes from seminal studies of the developing child’s ability to delay gratification (144). The earliest of these experiments were conducted in the late 1960s with child participants ranging in age from 4 to 6 years (and in other later experiments with kids all the way up through early adolescence). During the task, the child sits across the table from an experimenter who places a bowl of marshmallows, cookies, or some other tempting treat on the table between them. The child is told that the experimenter must leave the room for a few minutes. If the child can wait until the experimenter returns, she can have two treats, but if she can not wait, then she is allowed to have just one and must ring a bell (also located upon the table) to let the experimenter (who was in another room) know that this happened. The child is thus faced with a self-regulatory dilemma: to have one delectable treat now or to withhold desire for it in favor of having two treats later on. The idea here is that this choice between short- and long-term gains models for the developing child the kind of dilemmas adults face in everyday life, including choices like eating fattening foods and smoking cigarettes, whose immediate pleasures come at the price of poorer long-term health and longevity. In longitudinal studies, Mischel et al. found that the amount of time a child could wait to consume the treat predicts a number of important adult outcomes, including scores on standardized aptitude tests, income and education levels, and tendencies to have positive social relationships and not engage in substance abuse (145).

Recently, functional imaging studies have begun to examine the neural bases of this ability with a paradigm borrowed from behavioral economics known as temporal discounting. In the temporal discounting paradigm individuals are given a choice between receiving a smaller amount of money (or similarly valued consumer good) immediately as opposed to a larger sum of money (or more highly valued consumer good) at some time down the road (146). Individuals vary in the extent to which they’re willing to trade-off short-term cash-in-hand for a larger longer-term payoff, with some discounting the higher value of the longer-term gain to a greater extent than others. Imaging results (147) have shown that when individuals choose the immediate gain, activity is observed in regions associated with expressing preferences, affective learning, and reward (e.g., mPFC and ventral striatum). Strikingly, the loci of mPFC activation include ventral and perigenual regions similar to those implicated in outcome-based regulation. By contrast, when individuals choose the long-term gain, they show greater activity in dorsal and ventral lateral PFC as well as lateral OFC. Strikingly, these regions have all been implicated in description-based regulation as well as response selection and inhibition more generally. Thus (as outlined in Figure 5), regulating behavior through choice might involve a functional trade-off between systems involved in outcome-driven learning as opposed to guiding behavior on the basis of high-level mental representations of stimulus meaning (cf. 148). This might be because individuals might solve the delay dilemma in a variety of ways, including relying on their assessments of the current motivational value of a stimulus, which is updated as they pick the immediate gain, as opposed to using reappraisal, which might allow them to focus on the more abstract long-term goal (149,150).

In sum, important strides have been taken toward elucidating the neural bases of three ways of regulating behavior in a contextually appropriate manner. Nevertheless, a number of important questions remain. Perhaps foremost among them is the question of what specific computational processes are implemented in any putative control region and how that computation is recruited similarly or differently for each means of regulation. Another important question is how each type of regulation might differ as a function of the emotion or response one is attempting to control. The study of this topic is still quite new, and it remains to be seen why related but perhaps different regions are activated across studies of ostensibly similar forms of regulation.

Realizing the Translational Potential of the Framework

The goal of this review is to sketch a simple framework for organizing both basic and translational research on the neural bases of human social cognitive and emotional behavior. The juxtaposition of basic and translational approaches is important,
because translational research is always a two-step process. In the first step, basic research provides models for understanding normative behavior in healthy individuals. In the second step, translational research takes these models and applies them to clinical populations to help elucidate potential mechanisms of dysfunction. This two-step progression has been the model for prior cognitive neuroscience-inspired research on the neural bases of attentional and high-level cognitive deficits in schizophrenia (151–153). This article provides a blueprint for similar progress in the domain of social and emotional functioning.

The previous section of this article described the first step in this progression by providing a brief synthesis and synopsis of foundational controversies and debate about the fundamental neural bases at each ability/construct. Translational questions apply insights gained from basic research to understanding dysfunction in schizophrenia or other clinical disorders more generally. The gradient at the bottom of the Figure roughly represents our current state of basic knowledge about each ability/construct. Thus, the greatest amount of research has been devoted to understanding the first ability/construct and the least has been devoted to understanding ability/Constructs 4 and 5.

Construct 1: Translational Research on Acquisition of Social-Affective Values and Responses

Basic research has yet to fully investigate how the neural systems for affective learning might be involved in the acquisition of information about social as compared with non-social cues. As a consequence, translational research on that topic will have to wait, at least for the moment. Translational research could progress immediately, however, by building on one of the strongest foundations of research in all of social cognitive and affective neuroscience. Consider that behavioral research has begun to suggest that individuals with schizophrenia might have a normal internal experience of emotion in the moment but that they fail to anticipate or expect that future events will elicit these emotions (11). As mentioned in the preceding section, this distinction between the anticipation/expectation and immediate experience (or consummation) of a stimulus has been related to the function of the central striatum, mPFC, and amygdala (159). With that work as a foundation, imaging studies could use well-studied reward learning paradigms to determine whether individuals with schizophrenia fail to recruit the ventral striatum during the anticipation of a rewarding stimulus and mPFC when it is experienced. In like fashion, fear conditioning paradigms could be used to determine whether individuals with schizophrenia show the normal acquisition of conditioned responses (which are essentially expectations of an aversive stimulus) mediated by the amygdala. Once basic research clarifies the way in which social cues (such as facial expression) might depend upon these circuits, additional studies could help clarify when and how individuals with schizophrenia effectively recruit the neural systems for learning the affective significance of social as compared to non-social stimuli. This would allow determination of whether deficits are stimulus-general or specific to social stimuli per se.

Some work has already borne out some of these predictions in animal and human behavioral research. For example, animal models have suggested that affective learning deficits might be found in schizophrenia (160,161), and human behavioral studies have shown deficits in some forms of affective and non-affective conditioning (162–164) and reward-related decision-making (165). Human imaging studies have just begun to examine appetitive forms of learning in schizophrenia. The results of a handful of initial studies are converging to suggest that individuals with schizophrenia might fail to recruit reward related regions, such as the ventral striatum (166–169). This work has shown that striatal activation during anticipation predicts various negative symptoms, such as anhedonia, which augers well for future work seeking to relate neural markers of affective learning to negative symptoms and functional outcomes in schizophrenia.

Construct 2: Translational Research on Recognizing and Responding to Social-Afffective Stimuli

Basic research has begun to unravel the question of whether brain systems support the recognition of social-affective stimuli by implementing expression-specific or process-specific computations. As this work continues to unfold, translational research can begin investigating the way in which judgments about these stimuli vary as a function of one’s clinical status. The idea here...
would be to use knowledge of the specific functional roles played by specific brain systems to test hypotheses about social-emotional recognition deficits for a given population. In the case of schizophrenia one could ask whether an individual's tendency to fear or avoid others is related to dysfunction in neural systems supporting the recognition of different types of social stimuli. For example, it is possible that individuals with schizophrenia would show abnormal activation of the amygdala during the perception of faces that are considered to be untrustworthy (44–46).

Research could also move beyond this distinction to ask questions about systems involved in conscious as compared with non-conscious stimulus perception. In healthy individuals, the amygdala responds to untrustworthy faces even when they are not attended and responds to fear expressions even when they are presented subliminally (45,46,170). This has been taken as evidence for the relatively automatic encoding of such threat-related stimuli by the amygdala, and one could ask whether the automatic processing of these stimuli is disrupted in schizophrenia. If the automatic recognition of social-emotional cues is intact, then one could ask whether the conscious expression of evaluative preferences for stimuli, which has been shown to depend upon the mPFC, might be abnormal.

Alternatively, subtle nonverbal social cues might be especially problematic for individuals with schizophrenia, especially cues that ambiguously convey the intentions of another person. Thus one might expect heightened amygdala activation to neutral faces or to patterns of eye gaze that normally are not considered threatening in healthy individuals. Finally, paradigms used to investigate the role of the STS in recognizing nonverbal cues (e.g., biological motion) or the role of mPFC in expressing evaluative preferences might also be used to test hypotheses about the kinds of social cues that are problematic for individuals with schizophrenia.

In the past decade a great deal of work has been devoted to making progress on these issues by evaluating some kinds of social and emotional recognition in individuals with schizophrenia. In general, the results of this work support the idea that the mechanisms underlying Construct 2 might be dysfunctional in schizophrenia, with the majority of work focusing on their impairments in recognizing facial expressions of emotion (171) that might be persistent and predict functional outcomes (172,173), and that might be related to a broader deficit in recognizing faces in general, regardless of their expression (174,175). Functional magnetic resonance imaging and electrophysiological studies have begun to suggest that this deficit might involve reduced activity in the amygdala, insula, and related structures (176–179) as well impaired structural encoding of faces (180,181).

Construct 3: Translational Research on Embodied Simulation or Low-Level Mental State Inference

Perhaps the most important question facing basic research on embodied simulation/low-level mental state inference is whether and how activity in putative shared representation systems is related to actual behavior. Importantly, this includes behavioral measures of the ability to accurately identify or mimic the actions, thoughts, and emotions of others. As mentioned earlier, insights into these issues are just now appearing on the research horizon (100). As they come closer, translational imaging studies could turn to investigating three kinds of questions about the function of shared representations in individuals with schizophrenia. First, they could ask whether systems related to action programming, pain, or emotion are activated normally when they are experienced in the first person. Second, they could ask whether these systems are activated normally during the third-person observation of another person having these experiences. And third, they could ask whether the systems activated for first- and third-person action/experience overlap in individuals with schizophrenia in the same way and to the same extent that they do in healthy individuals, thereby providing evidence for the status of shared representations. The answers to these questions could have important implications for understanding social behavior in individuals with schizophrenia. For example, if patients fail to normally activate shared representations when observing the actions and experiences of others, they might lack some of the elements essential for building a direct experiential understanding of the internal states of others—an understanding that motivates prosocial behavior, helping, and the formation of social bonds (16,87).

Because research on this construct has thus far depended on the use of measures of overlapping brain activity, little behavioral work has explored potential deficits in schizophrenia. One notable exception is an electromyographic study showing normal emotion-related facial expressivity as well as expected facial mimicry responses to pictures of facial expressions (182). This might suggest that shared representations are intact in schizophrenia, at least to some degree. The task for future work will be to supplement behavioral studies of this construct, which themselves are relatively new, with imaging work examining these issues in individuals with schizophrenia. Other clinical populations (such as individuals with autism) suffering from impairments in social and emotional abilities have shown abnormal activity in shared representation systems (183), however, which suggests that imaging methods might be able to detect potential deficits in individuals with schizophrenia as well.
might infer that the semantic but not perceptual or inferential components of mental state inference have been impacted.

To date, behavioral and neuroscience research has made significant progress toward documenting deficits in the ability of individuals with schizophrenia to make mental state attributions. Notably, however, this work has not proceeded in the order suggested in the preceding text. For the most part, it has paralleled work on Construct 2, which concerns the bottom-up recognition of social-emotional cues in others, by examining the use of higher-level processes to make mental state attributions about those cues. This means that work has not yet carefully examined the extent to which the neural systems for attributions about one’s self and others are or are not common or distinct and are or are not impaired, in patients as compared with controls. Instead, as shown in recent reviews and meta-analyses, behavioral studies have focused on documenting consistent deficits in a variety of tasks requiring “mentalizing” about others as well as showing that these deficits might relate to actual social behavior and remain significant even after controlling for generalized cognitive deficits (184–186). In like fashion, imaging studies have begun investigating the neural correlates of performing these tasks and have found both functional deficits in activation of mPFC, amygdala, STS, and other components of the “mentalizing” network (9,187). To the extent that self-attributions about mental states have been examined, it has been with tasks examining perceptions of agentic control over action rather than attributions about emotions or traits (9). Future work might serve to bridge the gap between these literatures, perhaps in the ways suggested earlier.

Construct 5: Translational Research on Context-Sensitive Regulation

Although the neural systems involved in context-sensitive regulation of behavior have (to date) received the least attention (at least in human research) of any of the constructs, there is sufficient coherence in the extant data to provide the basis for translational endeavors (21). The primary emphasis of current basic science work has been on identifying the neural systems supporting the regulation of negative affective responses with reappraisal or extinction, and these methods could be extended to study the ability of individuals with schizophrenia to successfully recruit prefrontal control regions on the one hand and modulate systems involved in generating affective responses on the other. To the extent that individuals with schizophrenia have generally heightened tendencies to perceive threats and/or maintain top-down goals, regulation might prove difficult. This difficulty might manifest itself as heightened activity in the amygdala (or related structures) or diminished activity in medial or lateral PFC or both.

Few basic science studies have investigated the systems important for choice-based regulation, but the results do suggest competing hypotheses for schizophrenia. On one hand, patients might favor short-term gains to the extent that they are unable to recruit lateral prefrontal regions to maintain cognitive representations of long-term goals that can be used to inhibit affective responses to immediately available stimuli. On the other hand, patients would favor long-term gains to the extent that currently available stimuli generate no expectation for pleasure that competes (see Construct 2). Critical here will be the extent to which progress is made on the basic science front to determine whether different regulatory dynamics are involved for positive as compared with negative emotions and for different types of regulatory strategies (21). As these issues become clear, it might be possible to determine whether individuals with schizophrenia have problems not just with one of the three broad types of regulation described here but rather with specific ways of implementing reappraisal, with particular kinds of choices (e.g., between relatively rewarding as compared with relatively aver- sive options), or with extinction for particular kinds of affective responses.

Because basic science research on the neural bases of context-sensitive regulation have only begun to be established in the past few years, it is not surprising that translational work on this ability has only barely begun in individuals with schizophrenia and has moved ahead only a bit more in other populations, such as depressive subjects. Although the bulk of imaging work to date has focused on cognitive forms of regulation, the one study to directly study emotion regulation in individuals with schizophrenia examined the ability to behaviorally regulate emotion expression, which has been related to prefrontal activity in only one imaging study thus far (129). This behavioral study reported that individuals with schizophrenia might be impaired in the ability to upregulate but not downregulate the behavioral expression of positive emotion (188), an ability that might predict long-term mental health outcomes (189–191). Other clinical groups, such as depressive subjects, have shown apparent dysfunction in the prefrontal-amygdala dynamics underlying the successful use of cognitive strategies (such as reappraisal) to regulate emotion (192,193). It remains for future work to determine whether and how the neural bases of these and other forms of regulation are intact in schizophrenia or other disorders.

Conclusions

It is said that the purpose of science is to carve nature at its joints. For this article, the hope is that the current framework carves the biggest joints appropriately, even if it gets some of the smaller ones wrong. To a certain extent this is to be expected, given that social cognitive and affective neuroscience are disciplines that have come into their own only in the past 5–10 years. Indeed, it takes time for a field to mature and for core findings to become solidified. This consideration motivated the gradient shown at the bottom of Figure 6, which is meant to convey that (to date) basic research has provided the greatest breadth and depth of core findings for the constructs described at the left side of the figure, with Construct 2—which concerns the bottom-up recognition of social and emotional cues—being the most ready for immediate translational work. The exception to this rule is Construct 4, which concerns the use of top-down processes to draw inferences about mental states and traits. As noted earlier, although there is more basic work to be done, paradigms for tapping the core systems underlying this construct are sufficiently developed to provide reliable vehicles for translational research.

In large part, the need for more basic work in humans can be traced to the strengths and limitations of prior work that has been based primarily on animal models. Consider that the basic mechanisms underlying Constructs 1 and 2 (subsuming simple forms of affective learning and recognition) are conserved across species. For decades, this has meant that one could study them in a rodent or nonhuman primate model without the need of a technique like functional imaging to study them in humans. A problem arises, however, when one wants to move beyond these simple forms of learning to those subsumed under Constructs 3–5: the first two ability/constructs cannot account

www.sobp.org/journal
for social-affective abilities that depend upon higher-level processes present only in humans. Before the advent of functional imaging, it was difficult if not impossible to study the neural bases of abilities like mental state inference and certain forms of context-sensitive regulation (like reappraisal). In certain ways, this makes the rapid progress of basic research in the past decade all the more impressive and the prospects for translational research and the success of the CNTRICS initiative all the more exciting.

Completion of this article was supported by National Institutes of Health Grant MH076137 and National Institute on Drug Abuse Grant DA022541.

The author reports no biomedical financial interests or potential conflicts of interest.

I would like to thank Tor Wager for assistance in preparing Figure 2.


