Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex

Mauricio R Delgado, Jennifer S Beer, Lesley K Fellows, Scott A Huettel, Michael L Platt, Gregory J Quirk & Daniela Schiller

The ventromedial prefrontal cortex is attributed with various functions during valuation, affect regulation and social cognition. *Nature Neuroscience* asked a moderator to lead researchers in a dialogue on shared and distinct viewpoints of this region's roles.

The ventromedial prefrontal cortex is a unique cortical region with respect to its anatomical and functional heterogeneity. Commonly referred to as vmPFC, this structure broadly denotes a large portion of the ventral prefrontal cortex comprised of various subdivisions, each with its own intrinsic pattern of connectivity across the brain (Bhanji *et al.*, **Supplementary Note 1**). Perhaps not surprisingly, given anatomical classifications, vmPFC functionality has been attributed to various domains. Three of the more prominent perspectives of vmPFC function include: (i) the economic valuation perspective, which focuses on the representation of the value of real and hypothetical objects and actions to inform decision making (Lesley Fellows and Scott Huettel, **Supplementary Note 2**); (ii) the affective regulation perspective, which discusses how the vmPFC may be important for interpreting (and reinterpreting) affective information and flexibly altering responses according to situational demands (Daniela Schiller and Gregory Quirk, **Supplementary Note 3**); and (iii) the social cognition perspective, which highlights the vmPFC's role in judgments of oneself and other social agents, by simulating mental states to evaluating one's own and others' social behaviors (Jennifer S. Beer and Michael Platt, **Supplementary Note 4**).

Is there a predominant view with respect to vmPFC function? Or are there common elements across the three prominent perspectives that can help shape a more general definition and functional role for this brain region? This dialogue is presented to discuss the strengths of the arguments that support the economic valuation, affective regulation and social cognition perspectives, while highlighting the most important findings that inform and challenge each viewpoint. (For the full unabridged dialogue, see **Supplementary Note 5**.) Further, this dialogue aims to explain the conceptual overlap between these approaches and any scientific benefit that has been or can be gained by combining ideas from the distinct perspectives to better outline a functional role for the vmPFC.

Each pair of investigators representing a distinct perspective, covering both human and animal research, and presents one answer or argument for each of the discussion points below.

About your perspective

Mauricio Delgado (moderator): What is the functional role of the vmPFC as defined by each perspective? Are there specific anatomical definitions of vmPFC that have been 'functionally' defined within each perspective?

Lesley Fellows & Scott Huettel (economic valuation): There is a large body of functional MRI (fMRI) and electrophysiology evidence that the vmPFC carries information about the personal, subjective value of stimuli or actions (i.e., economic utility or motivational value). That is, signals have been identified in this region that scale with subjective value (for example, 'willingness to pay' to obtain a stimulus) across many classes of stimuli and actions. vmPFC activity is higher for options of greater value and lower for options of lesser value. The observation that neurons in vmPFC encode relative value independent of the specific eliciting stimuli has led to the proposal that vmPFC represents subjective value in a 'common currency', which in turn would support flexible decision making across contexts. There remains uncertainty about the anatomical specificity of this claim at the subregional level. Within vmPFC, functional distinctions seem to exist between lateral orbitofrontal cortex (OFC), medial OFC and subgenual medial prefrontal cortex (PFC). For example, neuroimaging data are consistent with the perspective that posterior subgenual vmPFC tracks the future value associated with a decision option whereas anterior vmPFC tracks experienced utility (for example, ref. 1). There exists considerable heterogeneity even across neuroimaging studies, however, and neither human lesion data nor primate electrophysiology data provide definitive evidence for or against this topography.

Daniela Schiller and Gregory Quirk (affective regulation): The vmPFC contributes to learning to inhibit maladaptive affective responses, regardless of valence. This is usually demonstrated via extinction of conditioned responses, whether they are aversive, appetitive or addiction-related. In rodents, vmPFC includes the infralimbic cortex (IL) and ventral half of prelimbic cortex. In humans, this region is anterior to and below the genu of corpus callosum, including Brodmann areas (BA) 11, 25 and ventral parts of 24 and 32. The subgenual anterior cingulate cortex has been specifically associated with extinction success in humans using fMRI[2] and has been proposed as the human homolog of the rodent IL. This region resides in BA 25, 24 and 32; however, other extinction

or inhibitory related activations, such as cognitive reappraisal and placebo effect, could extend to the more anterior BA 11[3].

Jennifer S. Beer and Michael Platt (social cognition): The social cognitive perspective defines the vmPFC, in the broadest sense, as including areas such as the orbitofrontal cortex, frontal pole, subgenual cingulate cortex and portions of the anterior cingulate gyrus: (BAs 11/12, 47, 25 and portions of 10, 24 and 32; **Supplementary Note 1**)[4]. The social cognitive perspective has been studied in both humans and nonhuman animals, but there may be a wider set of social cognitive functions that are supported by vmPFC in humans (or at least a wider set than can be investigated readily in animals). Its basic functional role from the social cognition perspective is to support evaluation and representation of the interpersonal qualities and minds of the self and of certain social targets. Human research suggests that one subdivision supports self-description as well as descriptions of close others and people perceived to be similar to the self (BA 10)[5,6]. Another subdivision appears to support motivated self-descriptions of self and social targets (BA 11)[7–9].

MD: What are some of the most important findings that support the proposed interpretation offered by your perspective?

LF & SH: The role of vmPFC in valuation is well-supported by converging evidence from the core methods of systems neuroscience. Human neuroimaging and primate electrophysiology studies show that vmPFC tracks the personal subjective value of a wide range of stimuli[10,11], during active decisions and even in the absence of choice[1]. Moreover, when specific features of a choice are most relevant to valuation (for example, healthfulness information for health-conscious individuals), vmPFC activity preferentially tracks those specific features[12]. There is also evidence that this signal is necessary for value-based decisions: vmPFC lesions lead to inconsistent preferences[13] and change how information is acquired and integrated to determine a choice option's value[14]. Such lesions also disrupt flexible value-based learning in humans and nonhuman primates[15,16].

DS & GQ: In rodents, interfering with vmPFC function during threat (fear) extinction training via lesion, pharmacological, stimulation and optogenetic methods impairs subsequent retrieval of extinction[17,18]. In addition to aversive conditioning, this is also true for appetitive Pavlovian conditioning[19], drug seeking[20,21] and avoidance[22,23]. In humans, neuroimaging studies have identified vmPFC activity during threat-memory modification using various strategies[24] such as extinction, reversal and cognitive emotion regulation, as well as avoidance learning[25]. Also, patients with post-traumatic stress disorder (PTSD) who show impairments in extinction learning and retrieval exhibit reduced vmPFC responses[26].

JSB & MP: In terms of BA 10 in humans, evidence comes from fMRI studies of the self-referent effect (i.e., superior memory for information encoded in relation to self). More specifically, studies comparing self-evaluations (of personal preferences or traits) to evaluations of other people robustly identify a region of vmPFC activation associated with self-evaluations as well as evaluations of other people to the extent that they are emotionally close to the self or are perceived to share traits with the self[6,27].

In terms of BA 11, evidence for vmPFC involvement in the motivated perceptions of self comes from fMRI studies and lesion studies examining the extent to which people evaluate themselves or others in ways that are self-protective, which are often more positive than warranted (for example, in relation to base rates, objective criteria or other people's judgment). These studies suggest that one functional network centered around the vmPFC (BA 11) mediates the depth of processing when evaluating the self, and another functional network centered around the vmPFC mediates self-evaluation that is most protective of self-esteem[7]. Finally, this region may play a role in 'loosening' decision thresholds to arrive at a desired evaluation[9]. However, in nonhuman animals, a confluence of evidence from lesion and single unit studies suggest that an area typically associated with anterior cingulate cortex (ACC) responds to the experiences and motivational states of others[4,28–30]. Importantly, single neurons in this area—unlike in OFC or dorsal ACC—respond selectively to rewards given to others or include the rewarding experiences of others within responses to rewards given to self[31].

MD: A follow-up question comes from LF & SH, who prompt each group to consider an observation (or set of findings) that does not fit as well within their particular perspective.

LF & SH: There are findings from the memory literature that are not so easy to fold into a value rubric. For example, work showing a role for vmPFC in memory integration[32], even for stimuli that are not obviously 'valuable'. If vmPFC played a more general role in integration, albeit with economic value a particularly important and common target, then the value perspective could be subsumed under a still larger framework.

Another challenging finding is that vmPFC damage does not necessarily affect the ability to make value ratings[33]. This could mean that the value perspective is incorrect—which would itself contradict the other evidence discussed in our commentary—or it could mean that value remains an underspecified construct. Much as William James said of attention, everyone knows what value is. Attention has turned out to reflect an array of more specific brain processes, and value may very well be as heterogeneous. Thus, the next key challenge will be specifying value more tightly and defining it formally.

DS & GQ: A recent study was able to induce a sense of panic in amygdala patients, raising questions about the role of amygdala in fear processing[34]. This is consistent with previous evidence for intact subjective experience of emotion in amygdala patients[35]. Together, these findings highlight the importance of dissociating basic threat-learning processes and their corresponding, albeit independent, subjective affective states. Another issue was raised by recent rodent studies reporting intact recall of extinction following optogenetic silencing of IL neurons[36,37]. Given that the same optogenetic manipulation during extinction training prevents subsequent recall, IL-vmPFC likely provides a 'teaching signal' for the formation of extinction memory outside of IL. But perhaps the most troubling is a set of findings showing permanent reduction of threat responses following extinction, as well as return of fear following manipulations presumed to affect the original fear association, blurring the distinction between new learning of safety association and unlearning (or reconsolidation blockade or update) of the original memory. These findings may require us to forgo the classic dichotomy between new learning and unlearning and consider a mixed model whereby affective modulation is determined by aggregating new trace formation with old trace reformation[38].

JSB & MP: It is not clear that there are troublesome findings if one is open to the idea that the vmPFC is not necessarily performing one process to account for all three perspectives. Within the human research on the social cognitive perspective, many studies don't replicate each other when researchers have delved past the simple self-versus-other contrast. For example, sometimes vmPFC has been associated with similarity between self and other, and sometimes it has been associated with dissimilarity. In this regard, more research is needed to precisely characterize how vmPFC mediates cognition about the

self and others, but its role in social evaluation (among social targets) is fairly robust.

Considering each other's perspectives

MD: Are there predictions regarding vmPFC function that might be better characterized by one perspective rather than another?

LF & SH: Because subjective value can be formally specified in models, the valuation perspective can, in principle, be dissociated from other perspectives on vmPFC function. The social perspective predicts domain-specificity in vmPFC function, whereas the valuation perspective predicts domain-independence. To date, the available evidence much better supports the idea of domain independence[39], but future work could identify selectively social processing[40]. The affective perspective predicts that vmPFC tracks the engagement of emotional processes like regulation, whereas the valuation perspective predicts that vmPFC tracks the value of stimuli subsequent to any regulation; again, the valuation predictions have been supported by recent research[41]. An early, influential view of vmPFC as important in response inhibition, based in part on observations that animals with vmPFC lesions perseverated after changes in the reward value of choices (i.e., in extinction or reversal learning), does not explain a large body of more recent evidence[42]. Loss-of-function methods have not been applied to distinguish between general value and specific social and/or emotional accounts of the functions of this region. However, there are human lesion studies supporting the claims that the ventromedial frontal lobe plays a role in value-based social decisions[43] and that the emotional response related to choice in value-based learning[44], observations that align better, in our view, with a valuation function that is engaged generally across economic, social and affective domains.

DS & GQ: The affective regulation perspective predicts enhanced vmPFC signaling whenever affective response inhibition is required, regardless of valence or social context. For example, during associative reward learning and reversal (where one stimulus predicts reward and another does not, and then these contingencies switch), inhibition of the no-longer-appropriate affective response to the stimulus that switched from rewarding to nonrewarding would increase vmPFC signaling despite the reduced value, whereas the valuation perspective would predict decreased vmPFC signaling when reward is removed. Both in fact have been observed[45], albeit in separate subregions within the vmPFC, where a more dorsal part tracked reward value and a more ventral part indicated inhibitory function.

JSB & MP: If a specific region within vmPFC performs a purely social function, then lesions or inactivation of that area should impair social functions but not nonsocial functions, such as economic decisions or affect regulation. However, if there are specific channels or cell-specific social computation channels separate from economic or affect regulation channels within this region, then gross damage to the area would have broader impairments. In that case, only molecular-genetic tools capable of silencing channels in a cell-specific manner could evaluate these hypotheses.

Existing literature already suggests there is at least some independence of the perspectives. Whereas evaluation and representation of social targets such as the self and close or similar others might be used in the service of modulating one's moods or assessing reward or value, these processes do not necessarily have to operate for social evaluation. The social cognition perspective suggests that vmPFC should show greatest recruitment when evaluating between social targets versus between nonsocial targets. Nonsocial targets might be just as likely to require affect regulation and value assessment (for example, food for dieters). Additionally, there is not strong evidence that reward and affect regulation processes are operating in a redundant manner with social cognition. For example, if vmPFC marks social information because of its reward value or because it requires affect regulation, one might expect vmPFC activation to increase in relation to the diagnosticity of self-versus-other information (i.e., value) or prejudiced responses to other groups (i.e., affect regulation or value). Yet these effects do not tend to be found in the literature[46,47]. Similarly, vmPFC responds equivalently to the value of social rewards and the value of monetary rewards in humans[1], but single-unit recordings within vmPFC and adjacent OFC have identified neurons specialized for encoding social rewards exclusively[48]. Such findings suggest there may be specialized channels for processing social information in medial PFC that are 'invisible' to fMRI.

MD: Consider a broad example of real-world interaction, for example, an employee receiving a reprimand from a manager. Could each perspective provide an account of vmPFC function during this event?

LF & SH: This experience can be framed in motivational terms: the reprimand would generate a negative value-signal tracked in vmPFC and in turn would influence both immediate behavior (i.e., social disengagement) and future behavior (i.e., learning from this negative feedback not to behave in the same way again). Patients with vmPFC damage are slower to learn from feedback in experimental tasks where feedback shifts dynamically[15], suggesting that this region is critical for flexibly shifting behavior under conditions like this scenario. Studies of real world behavior in such patients could, in principle, provide more direct evidence. vmPFC damage can disrupt social behavior[49], but whether a straight line can be traced from neural mechanisms of value representation in vmPFC to real world behavior is, for the moment, highly uncertain. This is particularly the case because several nearby regions, often injured together with vmPFC, including lateral OFC and lateral PFC, appear to play important roles in social–emotional processes (for example, emotion recognition, regulation) in their own right.

DS & GQ: Here the manager is a (social) stimulus, which changes in value and leads to a change in the employee's behavioral response. All three perspectives would predict the resulting behavior and changes in vmPFC activity, but each may employ different vmPFC subpopulations or projection pathways. The affective regulation perspective in particular would regard the manager as the 'conditioned stimulus' and predict decreased vmPFC activation, as typically seen during threat learning.

JSB & MP: There are socially specific components to this interaction—namely social hierarchy, an intangible mind of another person—and nonsocial components, such as negative feedback and error-related computations. Both social and nonsocial components could in principle be identified in vmPFC. In particular, hierarchy-related information has been identified in mPFC, although fMRI studies in humans and monkeys place the locus of signal change in slightly different areas. Single neurons in monkey vmPFC and medial OFC also respond to images of hierarchically dominant monkeys[48]. Feedback and negative prediction errors are associated more clearly with dorsal ACC, in social or nonsocial situations[28,50]. BA 10/32 might be involved in trying to understand the meaning behind the face value of the manager's words as a function of how similar the manager is perceived to be to the self[27]. BA 11 would be involved in dismissing the feedback[8] or modifying one's behavior to incorporate the feedback into future performance[9] depending on which would be most self-protective.

MD: Are there commonalities between these distinct interpretations of this example? How might the same phenomena be described by different terms from each perspective?

LF & SH: In natural settings, the three perspectives may describe the same phenomenon at different levels, with 'social' describing a context for behavior, 'value' describing how specific stimuli motivate actions in that context and 'emotion' describing the evoked affective states. Accordingly, the same process may have effects at each level. For example, brain activation observed following emotion regulation could reflect (i) engagement of control processes to shape the affective response to a stimulus, (ii) the resulting change in the value of that stimulus or (iii) a specific social distancing strategy used to facilitate regulation.

DS & GQ: It might be helpful to consider a more elaborated example: following the negative interaction, the manager then provides neutral or positive feedback every day during the following week. In this case, all three perspectives would predict enhanced engagement of vmPFC in response to the manager, albeit using different terms referring to different aspects of the interaction: (i) value: assigning increasingly positive value to the boss; (ii) affective regulation: extinction or update of the initial negative association; and (iii) social: evaluating the boss's thoughts and intentions.

JSB & MP: In general, the commonalities will depend on how one defines a social agent and a social interaction. In this particular real world example, there is redundancy in the operation of social cognitive, value detection and affect regulation processes, and this is primarily because it involves the operation of self-relevant feedback from a potentially important other person. Specifically, vmPFC helps individuals detect whether the feedback is valuable and whether it requires affect regulation as well as aiding in modification of social cognition in the service of affect regulation.

However, there are real-world examples other than this one in which the social cognitive processes believed to be supported by vmPFC would not be redundant with value and affect regulation. For example, computing the value and needed affect regulation in relation to a sweet food or a financial incentive may not necessarily involve social cognitive processes. Additionally, if someone received feedback from a computer, would vmPFC be seen to predict the same types of behavior that it is expected to when feedback is delivered by a manager? In both cases the feedback is self-relevant, but the examples differ in terms of whether the feedback is delivered by a nonsocial versus a social agent. More specifically, the question is whether these types of examples would recruit vmPFC activity in the same manner, either in terms of patterns of activation within vmPFC or its covariation with neural activity in other brain regions. A big question is how we identify other agents as individuals with motivations, goals, intentions and minds. It is not clear that agency is an all-or-none phenomenon, either in the world or in the brain.

Additional accounts of vmPFC function

MD: This dialogue involves three predominant views of vmPFC function. However, there are certainly other domains in which vmPFC is featured. For instance, vmPFC is often discussed in the memory literature (for example, memory retrieval) and is involved in the default mode network (DMN). How do these additional accounts inform each perspective?

LF & SH: Understanding the relationship between the active processes supported by vmPFC and the DMN remains a challenge. Of note, the subregions of vmPFC most linked to the DMN only partially overlap with those most linked to valuation; in general, DMN-related activation tends to be more anterior within vmPFC. Other functions linked to vmPFC (like memory retrieval) can be considered both as processes in their own right and as evokers of value-related processes (for example, successful memory retrieval generates reward signals), and context-specific value information may need to be retrieved from memory and integrated with other information to support decision-making.

DS & GQ: Other domains involving vmPFC include: (i) controllability, in which behavioral control over a given stressor reduces deleterious effects of that and subsequent stressors and facilitates subsequent extinction in rodents[51] and humans[52]. This 'coping and resilience' function of IL[53] can be seen as an extension of vmPFC-mediated inhibition of maladaptive affective responses. (ii) Social buffering in rodents, in which being exposed to an aversive CS in the presence of another rat reduces the expression of conditioned fear[54]. IL is activated under these conditions, but it remains to be determined if social buffering is IL-dependent. Similarly, humans show reduced conditioned defensive responses following vicarious safety learning by observing another person undergoing extinction[55], but this process appears to diminish vmPFC function[56]. This line of research may bridge the three perspectives by clarifying whether social and valuation processes access the vmPFC inhibition system to modify behavior. (iii) vmPFC is part of the DMN, and the functions attributed to DMN (for example, self- and other-related processing, thinking about the past and the future, etc.) are relevant to all three perspectives. Within the DMN, different 'hubs' and subcircuits have been identified that may hold the key to distinguishing the different perspectives: perhaps the vmPFC reflects the overlap of connectivity in all three domains, whereas differences would clearly emerge when considering the wider network.

JSB & MP: The involvement of vmPFC in autobiographical memory retrieval does not appear to be wholly redundant. Specifically, Macrae *et al.*[5] found that vmPFC activation increases when encoding self-relevant information and when encoding information that is later remembered (versus forgotten) but does not show an interactive effect (for example, activation is not greater for self-relevant information that is later remembered versus information that is forgotten). In terms of the DMN, vmPFC is certainly associated with the DMN, along with the posterior cingulate cortex and temporoparietal junction (TPJ). Notably, the DMN is implicated in both nonsocial and social functions, though a unifying theory of the psychological function of DMN remains elusive.

MD: Is there potential overlap between DMN and the function of vmPFC posited in each particular perspective?

LF & SH: To the extent that the DMN tracks in-the-moment experiences (for example, attending to one's own body states), the observed activation in (anterior) vmPFC for experienced value could be consistent with DMN function. And, intriguingly, another region within the DMN, posterior cingulate cortex, is often reported to be coactive with vmPFC in tasks that involve estimate of subjective value (for example, delay discounting[57]). But the overlap is not complete: vmPFC is typically only coactive with TPJ when a decision-making task contains a social component. So, while we agree that understanding the role of vmPFC within the DMN (and connectivity therein) will be a profitable direction for future research, we do not expect

that the vmPFC will end up as simply a subcomponent of the DMN.

DS & GQ: To our knowledge there hasn't been any systematic analysis describing the function of DMN in affective regulation. Nevertheless, reported regions of activation largely overlap with DMN hubs, but the exact overlap as well as whether changes in system-wide functional connectivity during rest and task in relation to regulation success should be examined. Since emotion is linked to bodily states and therefore internal states, we expect DMN activation to closely match vmPFC function. However, since vmPFC integrates information from multiple regions, it is possible that different coactivation patterns will emerge within different modalities, which may distinguish the three perspectives. Within the affective domain, vmPFC and other DMN regions tend to show task-related activation (rather than the typical deactivation) to learned safety stimuli, possibly dissociating fear and safety internal states or even separate safety contexts.

JSB & MP: In many ways, the answer to this question reflects how we should conceptualize the connection between the physiological properties of the DMN and any psychological function. Our reading of the literature is that DMN-related activation in vmPFC is pretty widespread and that other nodes in the DMN (for example, posterior cingulate cortex and TPJ) are also modulated by economic, regulatory and social tasks. Moreover, the DMN is defined based on resting state, but it remains unclear what it means to have a mind at rest, whether this is an important state to understand and how it relates to the manner in which it is operationalized in most studies (for example, looking at a fixation cross). This raises the corollary that we also need to distinguish DMN activations from resting state activity or deactivations.

Methodological and experimental considerations

MD: Are there methodological challenges for the interpretation of vmPFC function within each perspective (for example, signal dropout in fMRI)? Are there examples where one perspective can benefit from methods and/or approaches more typically used by research in another perspective?

LF & SH: All of the major methods of systems neuroscience have contributed to understanding value processing in vmPFC (for example, lesion studies in humans and nonhuman primates, fMRI, primate and rodent electrophysiology, among many). Close correspondence between human and animal work has been a particular strength of the existing literature, aided by the ease with which similar models can be used across species: for example, selecting the currently most valuable option, learning from reward and punishment. Most methodological challenges are inherent to specific techniques and not specific to the valuation perspective (for example, signal drop-out in fMRI, damage to nearby white matter tracts in lesions).

DS & GQ: In rodents, optogenetic studies confirm the affective regulation perspective for vmPFC mediation of fear extinction. However, there are two caveats: (i) IL output may only be needed during the initial learning of extinction[36,37] and (ii) the direct target of IL projections may not be the intercalated cells of the amygdala as previously thought, but rather the basolateral[58,59] or basomedial[60] amygdala. Also, IL receives discrete input from extinction-activated neurons in basolateral amygdala[61]. These data challenge the view of vmPFC as providing a 'safety signal' and instead suggest that IL output serves as a teaching signal instructing plasticity in the amygdala, thereby modulating the affective response[62]. This neuronal wiring scheme could inform the other perspectives, considering vmPFC's role as instructing changes in value and social processing in other regions. Furthermore, fMRI methods challenge all perspectives as attempts to fully understand what 'deactivation' in blood-oxygen-level dependent (BOLD) signaling reflects at the neuronal levels remain obscure.

JSB & MP: The research on humans from the social cognitive perspective relies heavily on fMRI studies, so signal dropout is a problem. The tricky part is the fact that BOLD signal multiplexes multiple physiological processes—including synaptic inputs reflected in local field potentials, local computations and outputs reflected in spiking, and vascular changes that reflect a host of processes[63–65]. This uncertainty makes a strong case for the need to apply multiple complementary methods, in both humans and animals. But of course that's tricky when asking questions about what may be uniquely human faculties. There are some lesion studies that have been conducted in humans, but naturally occurring lesions do not tend to neatly fall into the subdivisions identified by fMRI, so that approach is also not a silver bullet to solve this issue.

MD: What would be an example of an experimental test to distinguish the perspectives?

LF & SH: Probably the most straightforward tests would imbue social or affective stimuli with another sort of value that is (partially) uncorrelated with their social or affective qualities. In effect, researchers should evaluate whether multiple value signals could be observed—potentially in different regions within vmPFC or elsewhere—or whether vmPFC tracks a single, integrated value signal that incorporates multiple contributors. Lesions or other inactivation methods leading to double dissociations across tasks drawing on such putatively distinct signals would constitute perhaps the strongest evidence for these distinctions.

DS & GQ: While it may be difficult to fully orthogonalize the three perspectives of vmPFC function, a test of associative learning and update with self-relevant social feedback may provide separate sets of predictions for each perspective. Consider a protocol associating person A with providing self-relevant positive feedback (for example, 'you are an excellent employee') and person B with providing negative feedback (for example, 'you are a terrible employee') and then reversing these contingencies. During the acquisition phase, the valuation and social perspectives would both predict enhanced vmPFC signaling to person A (due to positive value and self-relevant feedback, respectively) but predict opposite patterns to person B (reduced vmPFC signaling due to negative value but enhanced signaling due to self-relevant feedback, respectively). The affective regulation perspective would predict no engagement (or deactivation) to both person A and B, as no inhibition is required at this stage. Following reversal of these contingencies, the affective regulation and social perspectives would now both predict enhanced signaling to person A, who switched from providing positive to negative feedback (requiring inhibition of positive affective response but still providing self-relevant feedback), whereas the valuation perspective would predict decreased vmPFC activity with the diminished positive value. All perspectives would now predict enhanced vmPFC signaling to person B (who increased in value), which requires inhibition of negative affective response and provides self-relevant feedback). All together, the predictions for each person at each stage throughout the task consist of three distinct patterns dissociating the three perspectives. Thus, reversal learning with self-relevant and valence-specific associative learning delivered by social targets may be an effective tool to dissociate the three perspectives

within the same experimental design. The same experiment could be repeated with nonsocial stimuli to see if this alters the pattern or location of vmPFC activation, and the feedback could vary from purely social to monetary.

JSB & MP: It may be possible to define tasks that partially orthogonalize or at least partially decorrelate these functions. Nevertheless, we remain deeply suspicious that any test could somehow completely cleave value and affect from any social interaction, thus making it difficult to fully disentangle these perspectives. Whereas social interactions may vary in their degree of value or affect, a social interaction devoid of value or affect seems very unnatural indeed. Even if such a paradigm could be conceived, it still doesn't address the issue of what pattern of results across these factors would support the idea that vmPFC operates in a psychologically similar fashion. These issues emphasize the importance of using complementary methods to test ideas about brain function. Functional manipulations of vmPFC (either activating it electrically or optogenetically, or inactivating or lesioning it) could determine whether changing the function of this area generate domain-specific alterations in behavior or cognition.

Toward a unified view of vmPFC function and future directions

Below we present excerpts from a larger discussion on 'value' as a potential core function of vmPFC, featuring a question from JSB and MP on whether the concept of value clearly extends across perspectives.

LF & SH: A challenge is to determine the generality of this value signal: does it influence social interactions and emotional functions rather generally, or is it more narrowly involved in decision-making or more narrowly still in specific kinds of decisions? The first possibility would potentially explain data, reviewed by others here, that suggests this region is engaged in various social, self-referential and emotional contexts that do not necessarily explicitly require decisions. We contend that valuation is the current best candidate for a core function of vmPFC.

JSB & MP: Value has been raised as the most likely candidate as a core function of vmPFC. Value certainly has the broadest implications, particularly when value is not conceptualized as valence. For example, it can be just as valuable to regulate behavior so that one approaches good things as it is to regulate behavior to avoid bad things. However, if value is the psychological function underlying results from all three perspectives, why has it been so difficult to find vmPFC activity associated with value in social evaluation? For example, vmPFC activity is associated with uncertain impressions of other people (in comparison to relatively more certain impressions)[66], and evidence that vmPFC is related to diagnosticity of person information is not robustly found across studies[67].

LF & SH: There is evidence that vmPFC lesions disrupt social evaluation, discussed above. Yet, more direct comparison will require experiments that independently manipulate social value and economic value, which poses challenges for operationalizing social value in a manner that does not simply reify economic utility. Moreover, social stimuli provide a particularly rich, complex and idiosyncratic source of value information. The processes that connect social information to value may thus differ dramatically across experimental contexts and across individuals, making it difficult to cleanly extract a social value signal without attaching economic value through the task design.

DS & GQ: Perhaps the difficulty in identifying value in social evaluation is related to difficulty in defining what social value is. Is there a pure construct of 'social value' that is clearly separable from affective regulation and economic value? For example, social value could be a byproduct of social consequences that indicates an outcome value (for example, increased social status leading to more power, social shaming leading to harm, etc.). It could also represent an affective response triggered by social evaluation (such as the disgust response driven by immoral behavior, anger at interfering behavior, fear of social repercussions, etc.). Another option is that value-related vmPFC activation would become apparent when another person influences economic or other types of value (for example, a person that is generous or trustworthy in economic exchange; a person cooking highly nutritious or tasty food, etc.). Isolating such cases may help defining a clear construct of social value and its relation to vmPFC function.

MD: What should the major future goals be for understanding vmPFC function in carrying out economic, emotional and social behavior? What are the next steps in neuroscience to integrate (or differentiate) these perspectives to better outline a functional role for the vmPFC?

LF & SH: We advocate for three major goals. First, the field needs clear tests that distinguish these different perspectives. Because researchers are generally 'process-oriented': they focus on one perspective and then see evidence for that perspective in the functioning of a brain region, yet if a task engages multiple regions (via multiple processes), the same results would also provide evidence for one or more other perspectives. Second, research should work toward better functional specification within vmPFC. To date, the topographical organization of vmPFC is much less understood than that of other brain regions (c.f. arguments for functional topography in lateral PFC). The apparently close homology in this region, anatomically and functionally, across humans and macaques could be usefully leveraged to this end, given the technical challenges of imaging this region at high resolution in humans. Third, the brain processes of value should be operationalized beyond that of simple economic value. vmPFC function will not necessarily map neatly onto economic concepts; indeed, an improved understanding of how the brain instantiates value computations could clarify thinking about decision making beyond neuroscience.

DS & GQ: Moving forward, our understanding of vmPFC could be enhanced by: (i) more studies that test predictions of different perspectives in the same experimental design; (i) more studies comparing appetitive and aversive conditioned stimuli in same animal or subject; (iii) moving from Pavlovian fear conditioning to more complex, decision-based, naturalistic paradigms (active avoidance, foraging or fear competition, social stimuli); (iv) identifying subregions and flow of information within the vmPFC (for example, dorsal to ventral, value-encoding preceding inhibition); (v) comparing vmPFC involvement to the adjacent dorsal mPFC involvement (prelimbic area in rodent) to identify mutual influence and hierarchical functional organization; and (vi) identifying critical targets of vmPFC (amygdala, ventral striatum, midbrain)[68] and network activation as a 'workspace' for different functions.

JSB & MP: One way to organize the existing literature is to more systematically investigate the functional networks associated with each perspective. All of these processes may recruit vmPFC function; for example, social and financial reward both recruit OFC[69], while affect regulation and self-evaluation involves vmPFC[70]. Techniques like multivariate pattern analysis (MVPA) and representational similarity analysis (RSA) will be helpful for understanding whether patterns of activation within subdivisions of vmPFC are similar

or different, and techniques such as psychophysiological interaction (PPI) analyses will be helpful for understanding how these processes elicit similar or different patterns of activity covariation between vmPFC and other brain regions. Complementary methods, in addition to fMRI, including lesions, *in vivo* physiology and application of cell-type-specific or activity-dependent optogenetic techniques, which are standard in rodent models, should be deployed in primates to better elucidate domain-specific modular networks. More sophisticated computational models will be crucial for teasing apart these possibilities and relating data across levels of analysis.

In the greater scheme of things, we must also be open to the possibility that vmPFC function does not neatly align with these psychological domains of processing, that vmPFC performs a broader set of computations that are employed in dealing with all these domains or that there are specific subcircuits within vmPFC that handle these domains. We must also be mindful that brains evolved to generate adaptive behavior within the contexts in which early organisms most often found themselves and that nonsocial contexts preceded social ones, inviting the possibility that social computations harness ancestral mechanisms that evolved solve the challenges of the physical environment[71].

Concluding statement

The following statement is a compilation of thoughts from all authors, prompted by a request from DS and GQ for a unifying framework (if any) for all three viewpoints.

The vmPFC is a commonly studied brain region, given its wide network of connections and broad influences on behavior. In this dialogue, researchers representing three distinct perspectives, each of which often highlights the vmPFC as a critical node, were tasked with discussing similarities, differences and challenges in interpreting vmPFC function from the viewpoints of economic valuation, affective regulation and social cognition. The ambitious (perhaps optimistic) goal of this piece was to propose a unifying framework across the perspectives aimed at informing future dialogs and research directions. Taking arguments from all perspectives, it is clear that the vmPFC's complex role in human behavior is supported by its connectivity with various regions (for example, striatum, amygdala, TPJ), which allows for the encoding of value, modulation of affective responses and processing of socially relevant information. Nevertheless, it is also clear that a potential unifying theory of vmPFC function is not yet possible because of important differences in how this region has been studied in each perspective.

First, while there may be shared computations within the vmPFC, it is difficult to clearly outline this function, given differences in experimental paradigms, definitions of key concepts and approaches across perspectives. This is nicely illustrated in the discussion points about 'value' possibly being a core function of vmPFC across modalities, but this argument may fall short when considering how value may be qualified in the social domain. Interesting proposals for future experiments that cut across the perspectives and special considerations of their limitations are briefly discussed during the dialogue.

Second, understanding the topographical organization of the vmPFC, which, as previously discussed by the economic valuation group, is not as well studied as other regions, and outlining functional networks that may co-opt distinct subsections of the vmPFC could provide a better roadmap to clarifying the common computations across perspectives. Indeed, understanding the vast scope of vmPFC inputs and outputs and its functional significance, as mentioned in the first discussion point, could prove useful for future investigations. This may facilitate testing of ideas, such as the affective regulation group's suggestion of BA 11 as a region that potentially links the three perspectives because various inhibitory strategies can involve social stimuli and can update of affective or subjective value (**Supplementary Note 5, Question 11**).

Third, there are methodological challenges to overcome, as discussed in previous discussion points. Despite the dialogue's participants' best efforts to represent each perspective broadly with respect to species, methodology and point of view, much of the existing data that inform the discussion rely on fMRI. As pointed out by the social cognition group, overlap of BOLD activations in fMRI studies could be misleading and obscure important functional differences that may be supported by networks of activation that may interact to support common processes across the three domains. As discussed in the final point about future directions, along with discussion among scientists representing different perspectives, advances in analyses and complementary methodology are necessary to fully characterize the contributions of the vmPFC and (perhaps more aptly) of the functional networks that co-opt subsections within the vmPFC to enable behaviors including but not limited to economic valuation, affective regulation and social cognition.

Note: Supplementary Information is available in the online version of the paper.

1. Smith, D.V. *et al. J. Neurosci.* **30**, 2490–2495 (2010).
2. Phelps, E.A., Delgado, M.R., Nearing, K.I. & LeDoux, J.E. *Neuron* **43**, 897–905 (2004).
3. Diekhof, E.K., Geier, K., Falkai, P. & Gruber, O. *Neuroimage* **58**, 275–285 (2011).
4. Apps, M.A., Rushworth, M.F. & Chang, S.W. *Neuron* **90**, 692–707 (2016).
5. Macrae, C.N., Moran, J.M., Heatherton, T.F., Banfield, J.F. & Kelley, W.M. *Cereb. Cortex* **14**, 647–654 (2004).
6. Murray, R.J., Schaer, M. & Debbané, M. *Neurosci. Biobehav. Rev.* **36**, 1043–1059 (2012).
7. Flagan, T. & Beer, J.S. *Front. Hum. Neurosci.* **7**, 450 (2013).
8. Hughes, B.L. & Beer, J.S. *J. Cogn. Neurosci.* **25**, 613–622 (2013).
9. Hughes, B.L. & Beer, J.S. *Neuroimage* **61**, 889–898 (2012).
10. Bartra, O., McGuire, J.T. & Kable, J.W. *Neuroimage* **76**, 412–427 (2013).
11. Abitbol, R. *et al. J. Neurosci.* **35**, 2308–2320 (2015).
12. Hare, T.A., Camerer, C.F. & Rangel, A. *Science* **324**, 646–648 (2009).
13. Henri-Bhargava, A., Simioni, A. & Fellows, L.K. *Neuropsychologia* **50**, 1536–1542 (2012).
14. Fellows, L.K. *Brain* **129**, 944–952 (2006).
15. Tsuchida, A., Doll, B.B. & Fellows, L.K. *J. Neurosci.* **30**, 16868–16875 (2010).
16. Noonan, M.P., Kolling, N., Walton, M.E. & Rushworth, M.F.S. *Eur. J. Neurosci.* **35**, 997–1010 (2012).
17. Sotres-Bayon, F., Cain, C.K. & LeDoux, J.E. *Biol. Psychiatry* **60**, 329–336 (2006).
18. Milad, M.R. & Quirk, G.J. *Annu. Rev. Psychol.* **63**, 129–151 (2012).
19. Rhodes, S.E. & Killcross, A.S. *Eur. J. Neurosci.* **25**, 2498–2503 (2007).
20. Peters, J., Kalivas, P.W. & Quirk, G.J. *Learn. Mem.* **16**, 279–288 (2009).
21. Van den Oever, M.C. *et al. J. Neurosci.* **33**, 18225–18233 (2013).
22. Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F. & Quirk, G.J. *J. Neurosci.* **34**, 9736–9742 (2014).
23. Fiorenza, N.G., Rosa, J., Izquierdo, I. & Myskiw, J.C. *Behav. Brain Res.* **232**, 210–216 (2012).
24. Schiller, D. & Delgado, M.R. *Trends Cogn. Sci.* **14**, 268–276 (2010).
25. Collins, K.A., Mendelsohn, A., Cain, C.K. & Schiller, D. *J. Neurosci.* **34**, 14733–14738 (2014).
26. Milad, M.R. *et al. Biol. Psychiatry* **66**, 1075–1082 (2009).
27. Mitchell, J.P., Banaji, M.R. & Macrae, C.N. *J. Cogn. Neurosci.* **17**, 1306–1315 (2005).
28. Chang, S.W. *et al. Proc. Natl. Acad. Sci. USA* **110 Suppl 2**, 10387–10394 (2013).
29. Rudebeck, P.H., Buckley, M.J., Walton, M.E. & Rushworth, M.F.S. *Science* **313**, 1310–1312 (2006).
30. Rushworth, M.F., Mars, R.B. & Sallet, *J. Curr. Opin. Neurobiol.* **23**, 436–442 (2013).
31. Chang, S.W., Gariépy, J.F. & Platt, M.L. *Nat. Neurosci.* **16**, 243–250 (2013).
32. Schlichting, M.L. & Preston, A.R. *Curr. Opin. Behav. Sci.* **1**, 1–8 (2015).
33. Vaidya, A.R. & Fellows, L.K. *Nat. Commun.* http://dx.doi.org/10.1038/ncomms10120 (2015).
34. Feinstein, J.S. *et al. Nat. Neurosci.* **16**, 270–272 (2013).
35. Anderson, A.K. & Phelps, E.A. *J. Cogn. Neurosci.* **14**, 709–720 (2002).
36. Bukalo, O. *et al. Sci. Adv.* **1**, e1500251 (2015).
37. Do-Monte, F.H., Manzano-Nieves, G., Quiñones-Laracuente, K., Ramos-Medina, L. & Quirk, G.J. *J. Neurosci.* **35**, 3607–3615 (2015).
38. Clem, R.L. & Schiller, D. *Trends Neurosci.* **39**, 340–351 (2016).
39. Levy, D.J. & Glimcher, P.W. *Curr. Opin. Neurobiol.* **22**, 1027–1038 (2012).
40. Ruff, C.C. & Fehr, E. *Nat. Rev. Neurosci.* **15**, 549–562 (2014).
41. Winecoff, A. *et al. J. Neurosci.* **33**, 11032–11039 (2013).

42. Stalnaker, T.A., Cooch, N.K. & Schoenbaum, G. *Nat. Neurosci.* **18**, 620–627 (2015).
43. Xia, C., Stolle, D., Gidengil, E. & Fellows, L.K. *J. Neurosci.* **35**, 8507–8514 (2015).
44. Bechara, A., Damasio, H., Tranel, D. & Damasio, A.R. *Science* **275**, 1293–1295 (1997).
45. Zhang, Z., Mendelsohn, A., Manson, K., Schiller, D. & Levy, I. *eNeuro* http://dx.doi.org/10.1523/eneuro.0072-15.2015 (2015).
46. Ito, T.A. & Bartholow, B.D. *Trends Cogn. Sci.* **13**, 524–531 (2009).
47. Mitchell, J.P., Cloutier, J., Banaji, M.R. & Macrae, C.N. *Soc. Cogn. Affect. Neurosci.* **1**, 49–55 (2006).
48. Watson, K.K. & Platt, M.L. *Curr. Biol.* **22**, 2268–2273 (2012).
49. Zald, D.H. & Andreotti, C. *Neuropsychologia* **48**, 3377–3391 (2010).
50. Hayden, B.Y., Pearson, J.M. & Platt, M.L. *Science* **324**, 948–950 (2009).
51. Amat, J. *et al. Nat. Neurosci.* **8**, 365–371 (2005).
52. Hartley, C.A., Gorun, A., Reddan, M.C., Ramirez, F. & Phelps, E.A. *Neurobiol. Learn. Mem.* **113**, 149–156 (2014).
53. Maier, S.F. & Watkins, L.R. *Brain Res.* **1355**, 52–60 (2010).
54. Kiyokawa, Y., Takeuchi, Y. & Mori, Y. Eur. *J. Neurosci.* **26**, 3606–3613 (2007).
55. Golkar, A., Selbing, I., Flygare, O., Ohman, A. & Olsson, A. *Psychol. Sci.* **24**, 2182–2190 (2013).
56. Golkar, A., Haaker, J., Selbing, I. & Olsson, A. *Soc. Cogn. Affect. Neurosci.* **11**, 1541–1549 (2016).
57. Kable, J.W. & Glimcher, P.W. *Nat. Neurosci.* **10**, 1625–1633 (2007).
58. Cho, J.H., Deisseroth, K. & Bolshakov, V.Y. *Neuron* **80**, 1491–1507 (2013).
59. Strobel, C., Marek, R., Gooch, H.M., Sullivan, R.K. & Sah, P. *Cell Reports* **10**, 1435–1442 (2015).
60. Adhikari, A. *et al. Nature* **527**, 179–185 (2015).
61. Senn, V. *et al. Neuron* **81**, 428–437 (2014).
62. Amano, T., Unal, C.T. & Paré, D. *Nat. Neurosci.* **13**, 489–494 (2010).
63. Logothetis, N.K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. *Nature* **412**, 150–157 (2001).
64. Logothetis, N.K. & Wandell, B.A. *Annu. Rev. Physiol.* **66**, 735–769 (2004).
65. Sirotin, Y.B. & Das, A. *Nature* **457**, 475–479 (2009).
66. Jenkins, A.C. & Mitchell, J.P. *Cereb. Cortex* **20**, 404–410 (2010).
67. Ames, D.L. & Fiske, S.T. *Neuroimage* **83**, 599–608 (2013).
68. Heilbronner, S.R., Rodriguez-Romaguera, J., Quirk, G.J., Groenewegen, H.J. & Haber, S.N. *Biol. Psychiatry* **80**, 509–521 (2016).
69. Lin, A., Adolphs, R. & Rangel, A. *Soc. Cogn. Affect. Neurosci.* **7**, 274–281 (2012).
70. Roy, M., Shohamy, D. & Wager, T.D. *Trends Cogn. Sci.* **16**, 147–156 (2012).
71. Pearson, J.M., Watson, K.K. & Platt, M.L. *Neuron* **82**, 950–965 (2014).

Mauricio R. Delgado is in the Department of Psychology, Rutgers University, Newark, New Jersey, USA. Jennifer S. Beer is in the Department of Psychology, University of Texas at Austin, Austin, Texas, USA. Lesley K. Fellows is in the Department of Neurology & Neurosurgery, McGill University, Montreal, Canada. Scott A. Huettel is in the Department of Psychology & Neuroscience, Duke University, Durham, North Carolina, USA. Michael Platt is in the Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania, USA. Gregory J. Quirk is in the Departments of Psychiatry and Anatomy & Neurobiology, University of Puerto Rico, San Juan, Puerto Rico, USA. Daniela Schiller is in the Departments of Psychiatry and Neuroscience and the Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA.