

Ventral anterior cingulate cortex in social decision-making

Patricia L. Lockwood^{1,2*} and Marco K. Wittmann^{1,2}

¹Department of Experimental Psychology, University of Oxford, OX1 3UD, Oxford, UK

²Wellcome Trust Centre for Integrative Neuroimaging (WIN), University of Oxford, UK

*Correspondence: patricia.lockwood@psy.ox.ac.uk (P. L. Lockwood), marco.k.wittmann@gmail.com (M. K. Wittmann)

Keywords: subgenual anterior cingulate cortex, perigenual anterior cingulate cortex, computational modeling, social decision-making, prediction error, self-esteem, empathy

Abstract: Studies in the field of social neuroscience have recently made use of computational models of decision-making to provide new insights into how we learn about the self and others during social interactions. Importantly, these studies have increasingly drawn attention to brain areas outside of classical cortical "social brain" regions that may be critical for social processing. In particular, two portions of the ventral anterior cingulate cortex (vACC), subgenual anterior cingulate cortex and perigenual cingulate cortex, have been linked to social and self learning signals, respectively. Here we discuss the emerging parallels between these studies. Uncovering the function of vACC during social interactions could provide important new avenues to understand social decision-making in health and disease.

1. Ventral anterior cingulate cortex during social interactions

Studies in the field of social neuroscience have focused on the neural mechanisms of how we understand and predict the actions, thoughts and feelings of other people (Joiner et al., 2017; Lee and Seo, 2016; Ruff and Fehr, 2014). While subcortical structures may be crucial for understanding valence-related aspects of the social environment such as recognizing emotional facial expressions (Adolphs, 2009), it is often the cortical regions of the brain that have been thought to underlie uniquely human social cognition and that have received comparatively more interest (Saxe, 2006). In particular the cortical network comprising the temporoparietal junction, precuneus and dorsomedial prefrontal cortex has been discussed extensively because of its putative role in inferring of others' mental states, often referred to as 'Theory of mind' (Frith and Frith, 2006; Gallagher and Frith, 2003; Saxe and Kanwisher, 2003; Schurz et al., 2014; Wittmann et al., 2018). The interest in this "social brain" network also increased in light of its anatomical overlap with the brain's default mode network (Mars et al., 2012).

More recently, however, it has become apparent that the medial prefrontal cortex (mPFC) might not have a unitary function in social cognition, but instead that there are several anatomical areas that comprise mPFC and that make

dissociable contributions. Prominently, in monkeys, lesions of the gyral part of anterior cingulate cortex (ACCg; areas 24a/b) decrease interest in conspecifics (Rudebeck et al., 2006) and in humans ACCg signals track information related to other people (Reviewed elsewhere, Apps et al., 2016, 2013; Behrens et al., 2008; Lockwood, 2016; Wittmann et al., 2018). It has been proposed that ACCg signals may reflect the motivational state of others (Apps et al., 2016; Lockwood, 2016) whereas the ACC sulcus dorsal to the gyrus may play a domain-general role in coding motivation of both self and others (Apps et al., 2016; Wittmann et al., 2018). In contrast to the ACCg and the above-mentioned dorsomedial parts of prefrontal cortex (Schurz et al., 2014), recent studies shed light on the role of more ventral portions of anterior cingulate cortex in social cognition, the perigenual anterior cingulate cortex (pgACC) and the subgenual anterior cingulate cortex (sgACC) (Palomero-Gallagher et al., 2015). In this mini-review, we highlight recent evidence for a role of sgACC and pgACC in how we learn about the self and others in social interactions.

Both sgACC and pgACC are part of agranular frontal cortex, which has correspondences in monkeys and also in rats (Wise, 2008). This is in contrast to granular dorsomedial areas involved in mental state inference, which appear to have evolved more recently (Wise, 2008). PgACC lies anterior to the genu of the corpus callosum and is cytoarchitectonically and functionally dissociable from sgACC, which occupies more ventral and posterior parts of cortex (Neubert et al., 2015; Palomero-Gallagher et al., 2015; Vogt et al., 1995, 1987). As expected from parts of cortex as close together as sgACC and pgACC, these areas have a similar connectivity profile (Neubert et al., 2015). Comparatively, however, sgACC is more strongly connected to many subcortical regions like the basal forebrain and hypothalamus than the pgACC, which in turn has relatively stronger connections to dorsal parts of medial and lateral frontal cortex (Neubert et al., 2015; Öngür et al., 1998; Price and Drevets, 2009). pgACC can also be distinguished from the midcingulate cortex in terms of functional and structural connectivity (Balsters et al., 2016b).

Very recently, studies have begun to uncover the role of these two vACC regions in social decision-making through the use of computational modelling

during social interactions (Diaconescu et al., 2017; Lockwood et al., 2016; Will et al., 2017; Wittmann et al., 2016). Whilst several models of decision-making exist, one particularly influential model comes from reinforcement/associative learning theory (RLT). At a basic level, RLT proposes that two key signals drive learning, a prediction error that measures the discrepancy between predicted and actual outcomes and an expected value signal that is updated by the prediction error (Rescorla and Wagner, 1972; Sutton and Barto, 1998). These models have been shown to be consistent with the firing rates of dopaminergic neurons in the midbrain (Schultz, 2013; Schultz et al., 1997) and several fMRI studies have shown prediction error and expected value signals that covary with responses in cortical and subcortical brain areas (see Chase et al., (2015) for a meta-analysis). RLT models have been used to study social cognition (Ruff & Fehr, 2014), often during so called strategic games like an iterated prisoners dilemma (Hampton, Bossaerts, & O'Doherty, 2008; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). These models and tasks have provided important insights into the role of ACCg for example in coding the changeability of other's behaviour (Apps et al., 2015; Behrens et al., 2008) and dorsomedial prefrontal cortex in coding other-directed expectation violations (Suzuki et al., 2012).

Seo and colleagues used RLT in combination with a competitive interactive game to identify cells in macaque medial frontal cortex that fire specifically when monkeys deviate from reinforced choice patterns, which is desirable in their paradigm because it enables to monkey to outplay the opponent (Seo et al., 2014). Often, the key interest of these types of studies of social interaction lies in the human or animal's ability for strategic reasoning given the observable behaviour of another player (Devaine et al., 2014). Albeit of considerable interest in itself, this focus on the strategic elements of social interactions cannot explain all aspects of real-life interactions (Schilbach et al., 2013). Specifically, there has been a significant and noteworthy absence of research on how we update our own thoughts and feelings about *ourselves* based on our interactions with others; accurately updating our beliefs based on interactions with others is critical for the formation of processes such as self-esteem and self-monitoring. Moreover, there has been a relative absence of

studies of the role of ventral, compared to dorsal ACC and adjacent mPFC areas in social decision-making.

Will and colleagues recently applied reinforcement learning models to characterise the neural mechanisms of self esteem (Will et al., 2017). They drew on principles from RLT to provide a novel psychological and computational account of how self-esteem develops over time, influenced by appraisals from others (Will et al., 2017). Participants received positive and negative feedback from other people that was allegedly related to an online profile of the participant set up before the study. Over time, participants learned to predict the other's evaluation of them. Every few trials, participants also rated how good they felt about themselves. Will et al (2017) found that participants felt more positive about themselves if they received positive feedback from others. Importantly, they felt this even more if that positive feedback was unexpected. In other words, a social prediction error - the difference between experienced and expected approval by others - shaped their self-esteem.

Will and colleagues were therefore able to relate self-esteem to the difference between predicted and received social feedback. Social prediction errors drove learning about others, but they also generated updates in one's own self-esteem. Intriguingly two different subdivisions of the anterior cingulate cortex tracked these social prediction errors and updates in self-esteem, respectively. Whilst the sgACC and adjacent ventral striatum tracked prediction errors in social feedback, activity in the pgACC reflected update signals (**Figure 1**).

2. Social prediction error signals in subgenual anterior cingulate cortex

The findings of Will et al dovetail with other recent work that has also identified prediction errors during social interactions in ventral portions of the medial prefrontal cortex (Diaconescu et al., 2017; Lockwood et al., 2016). These studies highlight sgACC and adjacent areas as critical. For example, Diaconescu and colleagues (2017) used a task where participants had to track an advisor's intentions over time, creating prediction errors when the advice was better or worse than expected. They found that the sgACC (areas 25 and

24s, (Palomero-Gallagher et al., 2015)) signaled prediction errors related to expected uncertainty about the advisor's trustworthiness (Diaconescu et al., 2017). Another study focused on prosocial learning, or how one's own actions affected outcomes for other people (Lockwood et al., 2016). They also identified social prediction error signals in sgACC (areas 25 and 24s, (Palomero-Gallagher et al., 2015)), but here these reflected whether the outcomes for another person based on our actions were better or worse than expected. Intriguingly this second study also found that sgACC signals were modulated by individual differences. In particular, those who reported themselves to be higher in empathy had greater tracking of social relative to self prediction errors in sgACC. Together these three new studies (Diaconescu et al., 2017; Lockwood et al., 2016; Will et al., 2017) suggest an important role for sgACC in social decision-making. Although several studies have linked sgACC function to prosocial and moral behaviours (Moll et al., 2005; Moll and Schulkin, 2009; Wiech et al., 2013; Zahn et al., 2009), only through the integration of computational fMRI have social subgenual signals relevant for decision-making begun to be uncovered.

Whilst the location of activation in these three studies (Diaconescu et al., 2017; Lockwood et al., 2016; Will et al., 2017) appears to be predominantly in the subgenual anterior cingulate cortex, particularly in areas 25 and 24s (as characterised by Palomero-Gallagher et al., (2015) and see Figure 1 for overlap) it is also important to consider adjacent regions to sgACC and the labels used to define this particular area. Posterior portions of the sgACC with the septal-anterior hypothalamic area that is part of the basal forebrain (Zaborszky et al., 2008) and can also extend into the ventral striatum (as in Will et al., 2017) and orbitofrontal cortex (OFC). However, the extent to which these areas play a domain-general or domain-specific role in social decision-making appears to differ (see Wittmann et al., (2018) for a review of ventral striatum and OFC in social behaviour). The ventral striatum tracked prediction errors in both the Diaconescu et al., (2017) and the Lockwood et al., (2016), but in a way that is suggestive of a domain-general response to learning about the unexpectedness of outcomes. For example, Lockwood et al., (2016) had a 'no-one' condition where points were delivered but participants were told they

would not be converted into money either for themselves or another person. Whereas ventral striatum tracked outcomes during self, prosocial and no one learning, sgACC specifically responded in the prosocial condition only. The OFC has also been linked to social cognition (Jones et al., 2011; Rushworth et al., 2007) but again seems more domain general than the sgACC. For example, Chang and colleagues (2013) recorded from OFC neurons in a reward-allocation task in macaques and found that these neurons predominantly responded to rewards delivered to oneself and not to others or no one. These findings support the idea that sgACC is relatively domain specific for social processing in certain contexts, but this hypothesis would need to be tested in further studies.

It should be noted that although the frame of reference in all three studies was related to the self, as the person in the scanner was interacting with another from the self perspective, there are perhaps interesting differences in the type of social PE signal in sgACC. In the studies by Will et al., (2017) and Diaconescu et al., (2017) the PEs were 'relational', they involved learning about others with clear consequences for the self, which impacted learning about self and other. In contrast in the Lockwood et al (2016) study the PEs occurred when learning about the outcomes for another person where the experimental outcomes for self were unaffected. It will be important for future studies to examine the commonality and distinction between different types of social PE in sgACC.

How do social computations in sgACC relate to sgACC's involvement in non-social processes? SgACC is closely connected to several areas involved in social cognition and reward processing (Moll et al., 2005). As mentioned above, compared to pgACC and other dorsal medial prefrontal areas, sgACC is more strongly connected to subcortical regions such as the basal forebrain than most areas in dorsal parts of medial prefrontal cortex (Beckmann et al., 2009; Moll et al., 2005; Palomero-Gallagher et al., 2015). These subcortical areas have long been linked to bonding, autonomic arousal and emotional responses (Moll et al., 2005; Rudebeck et al., 2014). Yet due to the difficulty in causing focal lesions to sgACC (Rudebeck et al., 2014), a paucity of studies in non-human

animals have been conducted. One recent study showed that lesions to the sgACC lead to disruptions in the maintenance of autonomic arousal associated with positive emotional events (Rudebeck et al., 2014). From this perspective, social prediction errors in sgACC might reflect the relevance of other people for one's own emotional state triggering changes in emotional arousal. On the other hand, studies of non-social decision-making have identified sgACC (area 25, (Palomero-Gallagher et al., 2015), Figure 1) responses to hierarchical prediction errors during sensory associative learning (Iglesias et al., 2013) linking this area to stimulus-outcome learning at levels of abstraction, perhaps similar to the type of abstract computations involved in real-life social interactions (Ruff and Fehr, 2014; Schilbach et al., 2013). However, sgACC is not traditionally characterised as part of the social brain, namely, areas that are preferentially recruited during social-cognitive processes.

3. Self-efficacy related computations in perigenual anterior cingulate cortex

If the sgACC computes prediction errors in social contexts what function distinguishes it from the pgACC? The most pertinent difference between sgACC and pgACC might be their frame of reference. Initial evidence suggests that whilst the sgACC might track information relevant to an other-centred frame of reference, the pgACC might compute information in a self-centred frame (Ruff and Fehr, 2014). In early investigations, for example, Kelley and colleagues found pgACC to be selectively active when judging whether adjectives describe oneself or not (Kelley et al., 2002). This was not the case when making attribute judgements related to other people.

More specifically, however, recent studies have discovered self-referential signals in pgACC that relate to what might be described as self-efficacy. Self-efficacy is a classical psychological concept that refers to the subjective belief that one will succeed in upcoming endeavours and overcome challenges (Bandura, 1977). Self-efficacy can have different sources. One way to increase self-efficacy is through positive feedback from other people. A role of pgACC in mediating this is in line with the finding from Will et al's (2017) that pgACC computes updates in self-esteem based on how positively oneself is evaluated

by other people (Will et al., 2017). PgACC indexes how approval by other people affects how good we feel about ourselves and our self-esteem is particularly boosted if we receive positive feedback from people that have evaluated us negatively in the past. Note that self-efficacy and self-esteem, although related, are clearly different psychological constructs (approximate correlation $r = 0.8$; Chen et al., (2004)) with distinct associations to being valued by others. While self-efficacy is more related to motivational variables and the belief to succeed in specific situations, self-esteem is rather related to emotional variables and a general self-worth.

A second source of self-efficacy is previous experience with a given task. If we have performed well in the past, we expect that we will also succeed in the future. A recent study investigated how humans learn about their past performance over the course of many trials (Wittmann et al., 2016). On every trial, subjects performed a short reaction-time based task and received parametric performance feedback for their own and another person's performance. Subjects were therefore able to build up a representation of their own ability in the task over time. Using trial-by-trial self-ratings, subjects indicated their expectation of how well they would perform on subsequent trials of the task. The authors found that subjects based their expectation of future success on their history of past performance. This was reflected in the BOLD signal in pgACC, that scaled with how good one's own task performance was in the past; higher levels of past performance were accompanied by higher BOLD activity in pgACC. In addition, the relationship between pgACC signal and past performance history was strongest in subjects who relied most on their past performance when predicting their future performance. This suggests that pgACC carries memory traces of past performance and that these guide expectations about how likely we are to succeed in future tasks. In contrast to dorsomedial prefrontal cortex, pgACC activity reflected relatively little information about the social context and the performance of other players, although it is important to note that pgACC activity occurred while the subjects were also concerned with tracking the performance of others in parallel.

In understanding the role of the pgACC in self and social decision-making it is also important to consider that pgACC is also often identified in non-social studies of value-guided decision-making. For example, pgACC reflects subjective value in delayed discounting experiments (Kable and Glimcher, 2007) where subjects have to trade-off the expectation of reward against a given timespan that they would have to wait before the reward is actually paid out. Similarly, pgACC also computes subjective value in some studies of self-control that require subjects to evaluate food items varying along the dimensions of tastiness (reward component) and healthiness (cost component) (Hare et al., 2011; Maier et al., 2015). Such a pattern of pgACC activity might reflect a general role in cost-benefit decision-making (Amemori and Graybiel, 2012) that is not specific to social or self-related processing during social interactions in particular. However, it is noteworthy that, first, subjective value and self-control in the aforementioned studies often partly depend on the current course of action, i.e. on estimates of one's upcoming actions, such as the ability to wait for future reward or to inhibit unhealthy food choices and that, second, more ventromedial prefrontal cortex carries representations of choice values that seem more pertinent to binary choice tasks than the value representations pgACC (Rushworth, Noonan, Boorman, Walton, & Behrens, 2011).

4. Implications for atypical social decision-making

More broadly, these new studies could have important clinical implications. Atypical social decision-making is commonly reported in several disorders including conduct disorder, psychopathy, autism and frontotemporal dementia (Anderson and Kiehl, 2012; Henry et al., 2016; Lockwood, 2016). Yet, very seldom have computational models been used to understand the neuroanatomical areas that are implicated. There is initial evidence that social prediction error signalling may be disrupted in those with autism spectrum disorder (Balsters et al., 2016a) and in social anxiety disorder (Koban et al., 2017). In addition, some studies have suggested different contributions of sgACC and pgACC to mood disorders such as depression (Drevets et al., 1997; Mayberg, 1997). The recent work reviewed here points to the utility of using

computational models along with self and caregiver report to provide new avenues to understand the neuroanatomy implicated in these disorders.

In terms of defining and diagnosing psychiatric and neurological conditions, there has been an increased focus on a transdiagnostic characterisation of different disorders that capture similar symptoms in different conditions (Huys et al., 2016; Montague et al., 2012). The study by Will et al (2017) utilises this relatively new approach to understanding associations between the brain and clinical conditions. They combine parameters from their computational model of self-esteem and self-report questionnaires of several symptoms including depression and anxiety. This allows them to identify an interpersonal vulnerability dimension across different conditions that modulated functional connectivity from medial prefrontal areas. Moving forward, this approach of harnessing well characterized frameworks of reinforcement and associative learning, along with transdiagnostic descriptions of psychiatric symptoms, promises to uncover new insights into clinically relevant psychological and neural mechanisms.

5. Summary

Overall, this article highlights an important emerging role for the ventral anterior cingulate cortex in social decision-making. Whereas the subgenual anterior cingulate cortex appears to track prediction errors during social interactions, the perigenual anterior cingulate cortex may reflect one's own estimates of future success. Moving forward, the application of computational frameworks of decision-making, close attention to neuroanatomy and learning models could provide a powerful way to understand disorders of social decision-making as well as how social decisions are made in everyday life.

Funding sources

This work was supported by an MRC Fellowship to P. L. L. (MR/P014097/1). The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).

341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367

References

- 368
369
370 Adolphs, R., 2009. The Social Brain: Neural Basis of Social Knowledge. *Annu. Rev.*
371 *Psychol.* 60, 693–716.
372 <https://doi.org/10.1146/annurev.psych.60.110707.163514>
373 Amemori, K., Graybiel, A.M., 2012. Localized microstimulation of primate
374 pregenual cingulate cortex induces negative decision-making. *Nat.*
375 *Neurosci.* 15, 776.

- Anderson, N.E., Kiehl, K.A., 2012. The psychopath magnetized: insights from brain imaging. *Trends Cogn. Sci.* 16, 52–60.
<https://doi.org/10.1016/j.tics.2011.11.008>
- Apps, M.A.J., Lesage, E., Ramnani, N., 2015. Vicarious reinforcement learning signals when instructing others. *J. Neurosci.* 35, 2904–2913.
<https://doi.org/10.1523/JNEUROSCI.3669-14.2015>
- Apps, M.A.J., Lockwood, P.L., Balsters, J.H., 2013. The role of the midcingulate cortex in monitoring others' decisions. *Front. Neurosci.* 7, 251.
<https://doi.org/10.3389/fnins.2013.00251>
- Apps, M.A.J., Rushworth, M.F.S., Chang, S.W.C., 2016. The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron* 90, 692–707. <https://doi.org/10.1016/j.neuron.2016.04.018>
- Balsters, J.H., Apps, M.A.J., Bolis, D., Lehner, R., Gallagher, L., Wenderoth, N., 2016a. Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain* 140, 235–246.
<https://doi.org/10.1093/brain/aww287>
- Balsters, J.H., Mantini, D., Apps, M.A.J., Eickhoff, S.B., Wenderoth, N., 2016b. Connectivity-based parcellation increases network detection sensitivity in resting state fMRI: An investigation into the cingulate cortex in autism. *NeuroImage Clin.* 11, 494–507.
<https://doi.org/10.1016/j.nicl.2016.03.016>
- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191.
- Beckmann, M., Johansen-Berg, H., Rushworth, M.F.S., 2009. Connectivity-Based Parcellation of Human Cingulate Cortex and Its Relation to Functional Specialization. *J. Neurosci.* 29, 1175–1190.
- Behrens, T.E.J., Hunt, L.T., Woolrich, M.W., Rushworth, M.F.S., 2008. Associative learning of social value. *Nature* 456, 245–249.
<https://doi.org/10.1038/nature07538>
- Chase, H.W., Kumar, P., Eickhoff, S.B., Dombrovski, A.Y., 2015. Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cogn. Affect. Behav. Neurosci.* 15, 435–459.
<https://doi.org/10.3758/s13415-015-0338-7>
- Chang, S. W. C., Gariépy, J.-F., & Platt, M. L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci*, 16(2), 243–50.
<http://doi.org/10.1038/nn.3287>
- Chen, G., Gully, S., Eden, D., 2004. General self-efficacy and self-esteem: toward theoretical and empirical distinction between correlated self-evaluations. *J. Organ. Behav.* 25, 375–395. <https://doi.org/10.1002/job.251>
- Devaine, M., Hollard, G., Daunizeau, J., 2014. The Social Bayesian Brain: Does Mentalizing Make a Difference When We Learn? *PLoS Comput. Biol.* 10, e1003992. <https://doi.org/10.1371/journal.pcbi.1003992>
- Diaconescu, A.O., Mathys, C., Weber, L.A.E., Kasper, L., Mauer, J., Stephan, K.E., 2017. Hierarchical prediction errors in midbrain and septum during social learning. *Soc. Cogn. Affect. Neurosci.* 12, 618–634.
<https://doi.org/10.1093/scan/nsw171>
- Drevets, W.C., Price, J.L., Simpson Jr, J.R., Todd, R.D., Reich, T., Vannier, M., Raichle, M.E., 1997. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 386, 824.

425 Frith, C.D., Frith, U., 2006. The neural basis of mentalizing. *Neuron* 50, 531–534.
 426 Gallagher, H.L., Frith, C.D., 2003. Functional imaging of ‘theory of mind.’ *Trends*
 427 *Cogn. Sci.* 7, 77–83.
 428 Hare, T.A., Malmaud, J., Rangel, A., 2011. Focusing Attention on the Health
 429 Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary
 430 Choice. *J. Neurosci.* 31, 11077. [https://doi.org/10.1523/JNEUROSCI.6383-](https://doi.org/10.1523/JNEUROSCI.6383-10.2011)
 431 [10.2011](https://doi.org/10.1523/JNEUROSCI.6383-10.2011)
 432 Henry, J.D., von Hippel, W., Molenberghs, P., Lee, T., Sachdev, P.S., 2016. Clinical
 433 assessment of social cognitive function in neurological disorders. *Nat.*
 434 *Rev. Neurol.* 12, 28–39. <https://doi.org/10.1038/nrneurol.2015.229>
 435 Huys, Q.J., Maia, T.V., Frank, M.J., 2016. Computational psychiatry as a bridge
 436 from neuroscience to clinical applications. *Nat. Neurosci.* 19, 404–413.
 437 Iglesias, S., Mathys, C., Brodersen, K.H., Kasper, L., Piccirelli, M., den Ouden,
 438 H.E.M., Stephan, K.E., 2013. Hierarchical prediction errors in midbrain and
 439 basal forebrain during sensory learning. *Neuron* 80, 519–530.
 440 <https://doi.org/10.1016/j.neuron.2013.09.009>
 441 Joiner, J., Piva, M., Turrin, C., Chang, S.W.C., 2017. Social learning through
 442 prediction error in the brain. *Npj Sci. Learn.* 2, 8.
 443 <https://doi.org/10.1038/s41539-017-0009-2>
 444 Jones, R.M., Somerville, L.H., Li, J., Ruberry, E.J., Libby, V., Glover, G., Voss, H.U.,
 445 Ballon, D.J., Casey, B.J., 2011. Behavioral and neural properties of social
 446 reinforcement learning. *J. Neurosci.* 31, 13039–45.
 447 <https://doi.org/10.1523/JNEUROSCI.2972-11.2011>
 448 Kable, J.W., Glimcher, P.W., 2007. The Neurobiology of Decision: Consensus and
 449 Controversy. *Neuron* 63, 733–745.
 450 <https://doi.org/10.1016/j.neuron.2009.09.003>
 451 Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F., 2002.
 452 Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14, 785–
 453 794.
 454 Koban, L., Schneider, R., Ashar, Y.K., Andrews-Hanna, J.R., Landy, L., Moscovitch,
 455 D.A., Wager, T.D., Arch, J.J., 2017. Social anxiety is characterized by biased
 456 learning about performance and the self. *Emotion* 17, 1144–1155.
 457 <https://doi.org/10.1037/emo0000296>
 458 Lee, D., Seo, H., 2016. Neural Basis of Strategic Decision Making. *Trends Neurosci.*
 459 39, 40–48. <https://doi.org/10.1016/j.tins.2015.11.002>
 460 Lockwood, P.L., 2016. The anatomy of empathy: Vicarious experience and
 461 disorders of social cognition. *Behav. Brain Res.* 311, 255–266.
 462 <https://doi.org/10.1016/j.bbr.2016.05.048>
 463 Lockwood, P.L., Apps, M.A.J., Valton, V., Viding, E., Roiser, J.P., 2016.
 464 Neurocomputational mechanisms of prosocial learning and links to
 465 empathy. *Proc. Natl. Acad. Sci.* 113, 9763–9768.
 466 <https://doi.org/10.1073/pnas.1603198113>
 467 Maier, S.U., Makwana, A.B., Hare, T.A., 2015. Acute Stress Impairs Self-Control in
 468 Goal-Directed Choice by Altering Multiple Functional Connections within
 469 the Brain’s Decision Circuits. *Neuron* 87, 621–631.
 470 <https://doi.org/10.1016/j.neuron.2015.07.005>
 471 Mars, R.B., Neubert, F.-X., Noonan, M.P., Sallet, J., Toni, I., Rushworth, M.F.S., 2012.
 472 On the relationship between the “default mode network” and the “social

473 brain." *Front. Hum. Neurosci.* 6, 189.
 474 <https://doi.org/10.3389/fnhum.2012.00189>
 475 Mayberg, H.S., 1997. Limbic-cortical dysregulation: a proposed model of
 476 depression. *J. Neuropsychiatry Clin. Neurosci.* 9, 471–481.
 477 <https://doi.org/10.1176/jnp.9.3.471>
 478 Moll, J., Schulkin, J., 2009. Social attachment and aversion in human moral
 479 cognition. *Neurosci. Biobehav. Rev.* 33, 456–465.
 480 Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., Grafman, J., 2005. The neural
 481 basis of human moral cognition. *Nat. Rev. Neurosci.* 6, 799.
 482 Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational
 483 psychiatry. *Trends Cogn. Sci.* 16, 72–80.
 484 <https://doi.org/10.1016/j.tics.2011.11.018>
 485 Neubert, F.-X., Mars, R.B., Sallet, J., Rushworth, M.F.S., 2015. Connectivity reveals
 486 relationship of brain areas for reward-guided learning and decision
 487 making in human and monkey frontal cortex. *Proc. Natl. Acad. Sci.*
 488 201410767.
 489 Öngür, D., An, X., Price, J.L., 1998. Prefrontal cortical projections to the
 490 hypothalamus in Macaque monkeys. *J. Comp. Neurol.* 401, 480–505.
 491 [https://doi.org/10.1002/\(SICI\)1096-9861\(19981130\)401:4<480::AID-](https://doi.org/10.1002/(SICI)1096-9861(19981130)401:4<480::AID-CNE4>3.0.CO;2-F)
 492 [CNE4>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1096-9861(19981130)401:4<480::AID-CNE4>3.0.CO;2-F)
 493 Palomero-Gallagher, N., Eickhoff, S.B., Hoffstaedter, F., Schleicher, A., Mohlberg,
 494 H., Vogt, B.A., Amunts, K., Zilles, K., 2015. Functional organization of
 495 human subgenual cortical areas: Relationship between architectonical
 496 segregation and connectional heterogeneity. *NeuroImage* 115, 177–190.
 497 <https://doi.org/10.1016/j.neuroimage.2015.04.053>
 498 Price, J.L., Drevets, W.C., 2009. Neurocircuitry of Mood Disorders.
 499 *Neuropsychopharmacology* 35, 192.
 500 Rescorla, R.A., Wagner, A.R., 1972. Classical Conditioning II: Current Research
 501 and Theory, in: Prokasy, W.F. (Ed.), . Appleton-Century Crofts, New York,
 502 pp. 64–99.
 503 Rudebeck, P.H., Buckley, M.J., Walton, M.E., Rushworth, M.F.S., 2006. A role for
 504 the macaque anterior cingulate gyrus in social valuation. *Science* 313,
 505 1310–1312.
 506 Rudebeck, P.H., Putnam, P.T., Daniels, T.E., Yang, T., Mitz, A.R., Rhodes, S.E.V.,
 507 Murray, E.A., 2014. A role for primate subgenual cingulate cortex in
 508 sustaining autonomic arousal. *Proc. Natl. Acad. Sci. U. S. A.* 111, 5391–
 509 5396. <https://doi.org/10.1073/pnas.1317695111>
 510 Ruff, C.C., Fehr, E., 2014. The neurobiology of rewards and values in social
 511 decision making. *Nat. Rev. Neurosci.* 15, 549–562.
 512 <https://doi.org/10.1038/nrn3776>
 513 Rushworth, M.F.S., Behrens, T.E.J., Rudebeck, P.H., Walton, M.E., 2007.
 514 Contrasting roles for cingulate and orbitofrontal cortex in decisions and
 515 social behaviour. *Trends Cogn. Sci.* 11, 168–176.
 516 <https://doi.org/10.1016/j.tics.2007.01.004>
 517 Saxe, R., 2006. Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16, 235–
 518 239.
 519 Saxe, R., Kanwisher, N., 2003. People thinking about thinking people - The role of
 520 the temporo-parietal junction in "theory of mind";
 521 *Neuroimage* 19, 1835–1842.

- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., Vogeley, K., 2013. Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. <https://doi.org/10.1017/S0140525X12000660>
- Schultz, W., 2013. Updating dopamine reward signals. *Curr. Opin. Neurobiol.* 23, 229–238. <http://dx.doi.org/10.1016/j.conb.2012.11.012>
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J., 2014. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* 42, 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>
- Seo, H., Cai, X., Donahue, C.H., Lee, D., 2014. Neural correlates of strategic reasoning during competitive games. *Science* 346, 340–343. <https://doi.org/10.1126/science.1256254>
- Sutton, R.S., Barto, A.G., 1998. Reinforcement learning: an introduction. MIT press, Cambridge, Massachusetts.
- Suzuki, S., Harasawa, N., Ueno, K., Gardner, J.L., Ichinohe, N., Haruno, M., Cheng, K., Nakahara, H., 2012. Learning to simulate others' decisions. *Neuron* 74, 1125–37. <https://doi.org/10.1016/j.neuron.2012.04.030>
- Vogt, B.A., Nimchinsky, E.A., Vogt, L.J., Hof, P.R., 1995. Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. *J. Comp. Neurol.* 359, 490–506. <https://doi.org/10.1002/cne.903590310>
- Vogt, B.A., Pandya, D.N., Rosene, D.L., 1987. CINGULATE CORTEX OF THE RHESUS-MONKEY .1. CYTOARCHITECTURE AND THALAMIC AFFERENTS. *J. Comp. Neurol.* 262, 256–270.
- Wiech, K., Kahane, G., Shackel, N., Farias, M., Savulescu, J., Tracey, I., 2013. Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. *Cognition* 126, 364–372.
- Will, G.-J., Rutledge, R.B., Moutoussis, M., Dolan, R.J., 2017. Neural and computational processes underlying dynamic changes in self-esteem. *eLife* 6, e28098. <https://doi.org/10.7554/eLife.28098>
- Wittmann, M., Kolling, N., Faber, N.S., Scholl, J., Nelissen, N., Rushworth, M.F.S., 2016. Self-Other Mergence in the Frontal Cortex during Cooperation and Competition. *Neuron* 91, 482–493. <https://doi.org/10.1016/j.neuron.2016.06.022>
- Wittmann, M.K., Lockwood, P.L., Rushworth, M.F.S., 2018. Neural Mechanisms of Social Cognition in Primates. *Annu. Rev. Neurosci.* <https://doi.org/10.1146/annurev-neuro-080317-061450>
- Zaborszky, L., Hoemke, L., Mohlberg, H., Schleicher, A., Amunts, K., Zilles, K., 2008. Stereotaxic probabilistic maps of the magnocellular cell groups in human basal forebrain. *NeuroImage* 42, 1127–1141. <https://doi.org/10.1016/j.neuroimage.2008.05.055>
- Zahn, R., de Oliveira-Souza, R., Bramati, I., Garrido, G., Moll, J., 2009. Subgenual cingulate activity reflects individual differences in empathic concern. *Neurosci. Lett.* 457, 107–110.

Figure legends

Figure 1. Self and social signals in perigenual anterior cingulate cortex (pgACC) and subgenual anterior cingulate cortex (sgACC). Table of studies that have identified responses in ventral anterior cingulate cortex during functional magnetic resonance imaging in humans using computational models of reinforcement learning. The anatomical locations from the peak co-ordinate in these studies are represented on the medial surface.