Citation in APA format: Decety, J., & Leong, Y. C. (2025). Social neuroscience. In D. T. Gilbert, S. T. Fiske, E. J. Finkel, & W. B. Mendes (Eds.), *The handbook of social psychology* (6th ed.). Situational Press. https://doi.org/10.70400/CACC8406

Social Neuroscience

Jean Decety, University of Chicago

Yuan Chang Leong, University of Chicago

WHAT IS SOCIAL NEUROSCIENCE?

Human beings are intensely social animals. Our lives intertwine with other people. Our health, economic prosperity, and well-being profoundly depend on our interactions with others. Social neuroscience seeks to understand how the brain mediates the dynamic relations between cognition and behavior in social contexts. This academic discipline is at the crossroads of biological sciences and social sciences. It relies on integrating knowledge from diverse disciplines and methods, on topics that are classically examined in social psychology, and this across multiple levels of analyses. Overall, social neuroscience aims to provide an overarching paradigm to investigate social behavior, and to determine where we as a species fit within a broader biological context (Cacioppo & Berntson, 1992).

The scope of social neuroscience is to specify the neural, hormonal, cellular, and genetic mechanisms and functions underlying social behaviors. The articulation across levels of analyses necessitates a vertical integration that requires that overlapping explanations at multiple levels of analysis be mutually compatible. Vertical integration does not impose that all scientists (be they economists, sociologists, psychologists, or biologists) say the same thing but that they come up with compatible and coherent explanations. It is important to note that vertically articulating different levels of organization is not synonymous with reducing or eliminating higher levels of analysis in favor of lower levels (Cacioppo & Decety, 2011). The theoretical constructs developed by social scientists (e.g., economists or sociologists) are as valuable as the constructs developed by biologists. However, analysis in the social sciences is often lateral, not vertical with little effort to explain phenomena by web of causation across adjacent levels of organization. Instead, social neuroscience is grounded in the premise that there are lawful relationships among biological, psychological, and social processes. It is thus essential to examine several levels to integrate their combined and dynamical properties.

This chapter provides a selective overview of this discipline by discussing contemporary theory and research on specific domains such as synchrony, social connections, social network, social hierarchies, dominance and status, motivated cognition, group dynamics, social influence on decision-making, and cooperation. It is not meant to be an exhaustive review of all psychological

dimensions that account for human sociability, but rather interrelated domains that represent some of the more current advances in social neuroscience, which also have direct and concrete applications in many aspects of human affairs, notably education, health, leadership, management and public policy. The chapter concludes with a discussion on some challenges and opportunities.

A Brief History Of The Discipline

The term "social neuroscience" was formally introduced in a 1992 *American Psychologist* article by John Cacioppo and Gary Berntson. The paper was intended to assess the future of human neuroscience and psychology and to facilitate the establishment of a disciplinary identity for academics working at the intersection of social psychology and neuroscience. Both disciplines had developed with weak interaction. Social psychology (with some exceptions) largely ignored biological sciences, whereas biological and social scientists came to realize that evolutionary theory could explain human social cognition and behavior (Alexander, 1974).

Social neuroscience emerged from social psychophysiology in the early 1990s as an interdisciplinary field devoted to understanding how biological systems implement social processes and behavior. It capitalizes on biological concepts and methods to inform and refine theories of social cognition and behavior, and using social and behavioral concepts and data to inform and refine theories of neural organization and function (Cacioppo & Berntson, 1992). These early pioneers were careful to note that there is no direct mapping of any physiological activity to a single psychological process and that social neuroscience is more than the study of the human brain (Vanman, Kappas, & Ito, 2021). Nearly a decade later, Cacioppo (2002) elaborated on his vision of social neuroscience, celebrating the collaboration of neuroscientists and cognitive scientists in the previous decade, and calling for more collaboration among social scientists and neuroscientists. When the fields combine, the resulting research is inevitably exciting and meaningful, not just to academics, but to the lay public as well.

The enthusiasm for this new discipline and its rapid growth led in 2006 of two new academic journals, *Social Neuroscience* and *Social Cognitive and Affective Neuroscience*, quickly followed by two scientific societies. The Social and Affective Neuroscience Society, founded in 2008, focused on the neural basis of human social and affective processes. Two years later, the Society for Social Neuroscience launched with, European, Chinese and Latin America Chapters and aims to integrate animal and human models of social behavior.

An Interdisciplinary Enterprise

Once acknowledged that social cognition is the product of reciprocal and dynamic influences between biological and social processes, it becomes obvious that a multidisciplinary perspective is crucial. Therefore, integrating theories and data from the social sciences and biological sciences significantly expands knowledge.

The interdisciplinary enterprise that characterizes social neuroscience capitalizes on a variety of disciplines, methods, and techniques. Some examples range from nationally representative surveys that include biomarkers and genetic data, behavioral studies, functional neuroimaging with electroencephalography or functional magnetic resonance imaging, pharmacological manipulations, neuroendocrinological and autonomic nervous system functions, hormones,

sociograms, and lesion studies with neurological patients. As an apt example, synchrony, the co-occurrence of 2 behaviors in time, is a widely spread phenomenon in many species including humans (Couzin, 2009). It can be examined with mathematical modeling of spatiotemporal firing patterns of neurons in vitro (Gritsun, Le Feber, Stegenga, &, Rutten, 2010), neural coupling in specific cortical and subcortical regions between a storyteller and their listener with fMRI (Yeshurun, Swanson, Simony, Chen, Lazaridi, Honey, & Hasson, 2017), synchronized neural activity in the frontal cortex in fruit bats when they socially interact (Rose, Styr, Schmid, Elie, & Yartsev, 2021), or in the physiological synchronization in cardiological interbeat intervals between group members at the individual-level and the group-level (Tomashin, Gordon, & Wallot, 2022). Such converging evidence suggests that a key emergent property of neural activity is synchrony across individuals, plays a pivotal role in the coordination of physiological processes, emotions, actions, mental representations, and resulting behavioral outcome such as cooperation. It further illustrates the added value of integrating multiple levels of analysis.

Multiple Levels Of Analyses

Given the complexity and variety of social interaction in humans, social neuroscience research needs to combine and integrate multiple-level analysis across different domains in order to determine the functional structures and computational processes that subserve social cognition and behavior (Cacioppo & Decety, 2009). Importantly, there is no one-to-one relation between psychological constructs and brain processes, particularly at the level of brain anatomy. Nor does there exist a simple one-to-one relationship between a structural MRI signal and the underlying brain structure. All the complex constructs that are examined in social neuroscience such as self-regulation, morality, empathy, trust, can hardly be seen as natural kind sited somewhere in the brain. They need to be broken down into sub-components, adding the challenge of avoiding narrow definitions that fail to capture the breadth of a function and vague concepts that may resonate with lay knowledge and common sense but have little, if any, scientific validity and utility. Molar constructs developed by behavioral and social scientists provide a means of understanding highly complex activity without needing to specify each individual action by its simplest components, thereby providing an efficient approach to describing complex systems (Cacioppo et al., 2007; Cacioppo, Berntson, & Decety, 2010). This is evident to most of us, but it needs to be regularly re-stated. For instance, one should not be surprised by a lack of an association between an anthropological theory of morality and regional gray matter volume (van Elk & Snoek, 2020). Moreover, because of the emergent properties of complex organizations, it is unlikely that an understanding of higher-level psychological processes can be satisfactorily derived from neuroanatomy and neurophysiology alone.

The explanatory power of integrative multilevel analysis is evident when attempting to model complex phenomena that develop over disparate levels of neural organization (Berntson & Norman, 2021). Multilevel analyses represent a subset of interdisciplinary approaches where the measures, constructs, and theories extend across levels of the organization and thus can be vertically integrated. Integrating across levels has many benefits. Consideration at one level can help generate novel hypotheses at the other, define categories of behavior, and set criteria that must be addressed in research (MacDougall-Shackleton, 2011). Vertical integration does not require that we all say the same thing, whether we are biologists, psychologists, sociologists, or economists, but that we say things that are compatible and mutually consistent. Even when holistic explanations are more appropriate and more coherent than reductionist explanations, they must not violate principles established at lower levels of explanation (Walsh, 1997).

Furthermore, interdisciplinary investigations of the commonalities and differences across social species are becoming increasingly important and successful. For instance, our understanding of empathy advanced when this phenomenological notion was broken down into distinctive dimensions such as affective, cognitive, and motivational. The affective dimension is consistent with emotional contagion that exists in a great many species, inducing social transmission of information, coordination, bonding, and shared experiences (Briefer, 2018; Nakahashi & Ohtsuki, 2015). In humans, for instance, the sharing of stories is driven in part by arousal. When people are physiologically aroused, whether due to emotional stimuli or otherwise, the autonomic nervous is activated, which then boosts social transmission (Berger, 2011). The cognitive dimension of empathy, also called theory-of-mind or perspective-taking, refers to the capacity to represent and reason about one's own and other people's intentions, emotions, and beliefs. Both affect sharing and theory of mind can induce a motivation to care for the other's welfare. These affective, cognitive, and motivational dimensions of empathy interact and rely on partly separable neural circuits that have specific mechanisms, which reflect evolved abilities and respond to social signals necessary for surviving, reproducing, understanding self and others, and maintaining well-being (Decety, 2015; Kanske, Böckler, Trautwein, & Singer, 2015). This is a mechanistic explanation why empathy facilitates prosocial behavior. Empirical research with humans (e.g., Lamm, Batson, & Decety, 2007; Lamm, Decety, & Singer, 2011) and rodents (Bartal, Decety, & Mason, 2011; Ben-Ami Bartal et al., 2016) demonstrates that empathy is implemented by a network of distributed, recursively connected, interacting neural circuits, including the brainstem, amygdala, hypothalamus, striatum, insula, anterior cingulate cortex, medial prefrontal cortex, temporoparietal junction, and orbitofrontal cortex, as well as autonomic nervous system and the neuroendocrine system (Decety, Norman, Berntson, & Cacioppo, 2012). Moreover, the motivation to care for and help others is associated with a feeling of satisfaction, a phenomenon known to both psychologists (Brown & Cialdini, 2015) and economists (Andreoni, 1990) alike. Neural activity in the reward-pathways and regions representing values is associated with helping behavior in rodents (Ben-Ami Bartal et al., 2021; Breton et al., 2022), and humans (Ashar et al., 2017; Decety & Porges, 2011), and also influence decisions to give money to people who are suffering (Genevsky, Västfjäll, Slovic, & Knutson, 2013). Across multiple studies, altruistic decisions to help or to give resources consistently activate the reward and value-computation networks (Cutler & Campbell-Meiklejohn, 2019). Evolutionary models of cooperation require proximate mechanisms that sustain prosociality despite inherent costs to individuals. The warm glow that often follows prosocial acts could provide one such mechanism.

Overall, the interdisciplinary perspective that characterizes social neuroscience articulates different levels of analysis in order to understand social cognition and behavior in a more comprehensive way, taking into account the inherent complexity of social phenomena. When interdisciplinary analyses integrate approaches from the social and biological sciences, they greatly expand our theoretical knowledge, both descriptive and predictive.

The Scope Of Social Neuroscience

Human biology has shaped the social environments we have created, and the social environment and cultures have impacted gene expressions, brains, and bodies. Feedback loops between humans and the environment have modified selection pressures on organisms and their descendants (Laland, Odling-Smee, & Myles, 2010). As a result of natural selection and gene-culture coevolution, several major features characterize our species (Table 1), and make our species unique (Boyd, 2017). Our ability and our need to be connected is unmatched with that of other primates, but we also

share a biological heritage with other species. Nonconscious influences are responsible for complex social behaviors that emerge in most if not all nonhuman social species, such as hierarchical organization and social dominance. Nonconscious mechanisms have been conserved in humans because they save energy and computational power, perhaps to a greater extent than is apparent from analyses of social cognition and behavior. Investigation of other social species, animal models of specific aspects of human functioning, and investigations of our evolutionary lineage provide important information regarding these conserved mechanisms. This then stimulates new theories and research on what might be unique to humans.

***Table 1*: Major Features Characterize Our Species**

- Reduced sexual dimorphism
- Neoteny imposed by bipedalism, which calls for a protective social organization and favors late cerebral development
- Community reproduction (alloparental assistance)
- Language
- Shared intentionality
- Importance of social information and communicating it, which allows regulating interactions with others
- Morality as a sense of obligation that humans feel toward one another
- Capacity for collective action
- Social emotions, such as envy, guilt, remorse, shame, and pride
- Receptivity to socialization and learning from others
- High levels of cooperation with non-kin that are enforced by social norms and control mechanisms such as punishment and reward, reputation, or ostracism
- Diffusion of sociality through social networks
- Gene-culture co-evolution

Social neuroscience has an important contribution in unraveling the mechanisms that connect us with our conspecifics and how social influences reciprocally interact in shaping these neural processes. The interdisciplinary perspective, which characterizes social neuroscience by integrating the biological and social sciences considerably broadens our knowledge and provides a more complete understanding of human sociality.

SYNCHRONY

Social interactions depend heavily on a shared understanding of the world (Frith & Frith, 2007). From a game of chess to having a conversation, joint activities are grounded in the shared understanding of the meaning of arbitrary rules and symbols. In infants, this shared representation is facilitated via mechanisms of joint attention - as early as 6 months of age, infants begin to coordinate attention with their parents through a combination of eye-gaze, gestures and utterances (Moore, Dunham, & Dunham, 2014). Infants also share emotional experiences with their caregivers via synchronization of physiological signals (e.g., heart rate, cortisol levels). For example, Waters and colleagues (2014) found increased mother-infant physiological synchrony after the mothers were subjected to a stressful experience. The synchronization of physiological responses may facilitate a sense of shared affective experience and help establish a foundation for more complex social competencies, including language development, empathy and perspective-taking. As a child

acquires language, shared representations can be transmitted via verbal or written communication.

The tendency to synchronize is one of the most pervasive drives in the biological world as illustrated by fireflies that not only flash in unison but in rhythm at a constant tempo. Spontaneous synchronization in people has been observed across a wide variety of actions and social settings – from the synchronized clapping at the end of a concert (Néda et al., 2000) to the synchronized gait patterns of people walking side by side (Nessler & Gilliland, 2009). In addition to overt behavior, affective states and physiological signals are also synchronized across individuals experiencing the same event (Cheong et al., 2020; Kang & Wheatley, 2017; Konvalinka et al., 2011). Interestingly, an individual's tendency to synchronize to a physical stimulus (e.g., the rhythm of a sequence of musical tones) has been found to correlate with their tendency to synchronize physiologically with another individual during a shared experience, suggesting that the tendency to synchronize might be a stable individual difference that generalizes across contexts (Wohltjen et al., 2023).

Synchrony Between Brains

How do two brains create a shared representation of an ongoing event or joint task? One mechanism by which this can be achieved is via the synchronization of neural representations across individuals (Nummenmaa et al., 2018; Wheatley et al., 2012). Over the past two decades, advances in neural recording techniques and analytical tools have made possible the study of the synchronization of neural signals that underlie synchrony in behavior and physiological responses.

Interbrain synchronization of neural activity has been studied across imaging modalities, including fMRI, EEG, ECoG, and fNIRS (Dieffenbach et al., 2021; Haufe et al., 2018). Researchers are thus able to take advantage of the strengths of each method to examine interbrain synchrony across different timescales, at different spatial resolutions and in varied environments. For example, fMRI is particularly useful for examining interbrain synchrony across the whole brain, including in subcortical areas that are difficult to reach using EEG or fNIRS (Sachs et al., 2020). However, the constraints of the MRI scanner limit experimenters to paradigms where participants lie in the scanner with limited capacity to interact with the outside world. On the other hand, fNIRS has limited spatial coverage, but is relatively inexpensive and portable, and can be used to study interbrain synchrony as participants engage in live conversation in naturalistic settings (Nguyen et al., 2021).

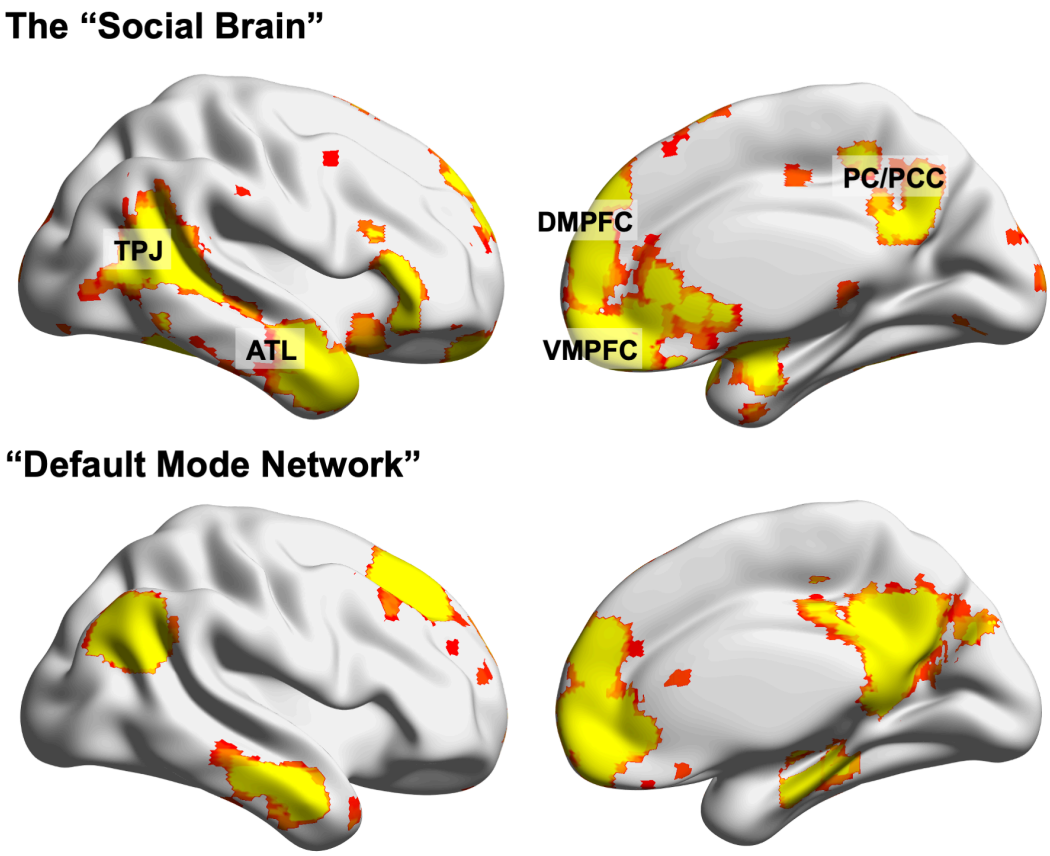
Interbrain synchrony can emerge when individuals' neural activity couples through perceiving the same sensory information. In an early fMRI study, the time course of neural activation was found to be synchronized across individuals watching the same movie (Hasson et al., 2004). Synchrony was quantified by computing the average Pearson correlation in brain activity over time across participants ("intersubject correlation" or ISC). ISC is thought to capture shared processing of a common stimulus (Nastase et al., 2019). In particular, the brain activity of two individuals not engaged in a shared experience is expected to be uncorrelated. Watching the same movie, however, drives brain activity in both brains. The correlation in activity across brains would thus reflect the extent to which a brain area is responding to a stimulus in a similar fashion across individuals.

Importantly, brain areas only synchronize across individuals if they are engaged by the shared stimulus. An fMRI experiment that provides an intuitive demonstration of this process measured the neural activity of Russian speakers listening to a Russian story, English speakers listening to the

same Russian story in Russian, and the same English speakers listening to an English translation of the same story (Honey et al., 2012). When comparing the neural activity of Russian and English speakers listening to the story in Russian, synchrony was observed in the auditory cortex, consistent with the shared processing of the acoustic properties of the stimulus. Since only the Russian speakers were able to understand the story, the semantic and narrative content were not shared between the two groups, and synchrony was not observed in higher order brain regions. When comparing the neural activity of Russian speakers listening to the story in Russian and English speakers listening to the English translation, synchrony weakened in the auditory cortex, consistent with the processing of a different auditory stimulus between groups. However, synchrony emerged in higher order brain regions such as the dorsomedial prefrontal cortex and precuneus, suggesting that these areas were involved in processing the semantic and narrative content of the stimulus, which were now shared between the two groups.

As such, assessing which brain regions synchronize under what conditions can reveal the functional role of a brain region, and can be used as a metric to determine the extent to which cognitive representations are shared across individuals. Using this approach, studies have identified a core set of brain regions that are more synchronized when subjects have similar interpretations of a narrative, including the dorsomedial prefrontal cortex, ventromedial prefrontal cortex, posterior cingulate/precuneus, temporoparietal junction, anterior temporal lobe (Nguyen et al., 2019; Yeshurun et al., 2017). These regions overlap with the "default mode network", which has been implicated in maintaining internally constructed representations (Buckner & DiNicola, 2019), as well as the network of regions thought to constitute "the social brain" (Meyer, 2019) (Figure 1). One proposal suggests that these regions are central in integrating internal (e.g., prior beliefs and goals) and external information (e.g., observed behaviors of other social agents) to maintain a neural representation of ongoing events (Yeshurun et al., 2021). Synchrony in these regions across brains then reflects shared representations between individuals.

***Figure 1*: The "Social Brain" and the "Default Mode Network"**



Note: The network of brain regions reliably engaged by social cognitive tasks (the "social brain") overlaps with the set of brain regions implicated in maintaining internally constructed representations (the "default mode network"). The brain maps were created by entering the terms "social" and "default" into Neurosynth (https://neurosynth.org/), a web-based platform for automated neuroimaging meta-analysis designed to identify brain regions that are most strongly associated with a term. DMPFC: dorsomedial prefrontal cortex; VMPFC: ventromedial prefrontal cortex; PC/PCC: posterior cingulate/precuneus; TPJ: temporoparietal junction; ATL: anterior temporal lobe. The maps were then visualized using BrainNet Viewer (Xia et al., 2013).

Factors That Promote Interbrain Synchrony

When do brains become synchronized and when do they fall "out of sync"? Below, we highlight three core factors that promote interbrain synchrony: (1) shared perspective, (2) joint attention and (3) communication.

Shared Perspective

In the 1950s Akira Kurosawa film *Rashomon*, the story of the murder of a samurai, is recounted by four eyewitnesses, each offering a contradictory account of what happened. In what has become known as the "Rashomon Effect," an identical series of events can be seen and interpreted differently by a set of seemingly reliable observers. Goals and beliefs have a powerful influence over one's view of reality. Unsurprisingly, having shared goals and beliefs are associated with stronger interbrain synchrony. Lahnakoski and colleagues (2014) had participants watch the same 10-minute movie clip twice: once assuming the perspective of a detective trying to solve a murder

depicted in the clip, and once assuming the perspective of an interior decorator tasked to redesign the set. Interbrain synchrony was stronger when participants shared a perspective than when they adopted different perspectives. Similarly, Yeshurun and colleagues (2017) manipulated participants to have different interpretations of an ambiguous story, and found that interbrain synchrony was stronger between participants manipulated to have the same interpretation. In both studies, researchers were able to accurately classify the perspective a participant adopted by assessing the level of interbrain synchrony with other participants.

To examine how interbrain synchrony is modulated by social context, Zhou and colleagues (2022) had Chinese participants watch separate documentaries about life as a high school student in the US and in China while undergoing fNIRS. One group of participants watched the documentaries with a US partner (US-primed group), while another group watched them with a Chinese partner (Chinese-primed group). During the documentary about US students, interbrain synchrony of US-primed participants was greater than that of Chinese-primed participants. In contrast, during the documentary about Chinese students, interbrain synchrony of the Chinese-primed participants was greater than that of US-primed participants. This priming effect was particularly strong during moments in the documentary that were culturally representative and was related to self-report measures of cultural perspective-taking. The authors interpret their results as suggesting that cultural cues prime individuals to adopt the corresponding cultural perspective, which promotes greater consistency in processing culturally-specific information.

Shared perspectives do not have to be explicitly manipulated, but can also arise from shared predispositions, beliefs or experiences. When viewing or listening to the same stimuli, interbrain synchrony is stronger between friends (Hyon et al., 2020; Parkinson et al., 2018), as well as between people who grew up in similar cultural backgrounds (Hakonen et al., 2022). While listening to a story where the intentions of the characters are ambiguous, individuals with high trait-level paranoia exhibited stronger interbrain synchrony than individuals with low trait-level paranoia, suggesting that paranoid individuals were predisposed to infer similar (likely nefarious) intentions of the characters (Finn et al., 2018).

Relatedly, researchers have also shown that partisans with similar political ideology exhibit strong interbrain synchrony while watching political content (e.g., news, campaign ads, political debates) (de Bruin et al., 2023; Dieffenbach et al., 2021; Katabi et al., 2023; Leong et al., 2020; van Baar et al., 2021). Interbrain synchrony between the politically like-minded was strongest while watching polarizing and emotionally engaging content, and predicted subsequent political attitudes. Individuals who were less tolerant to uncertainty in daily life experienced the strongest interbrain synchrony with those who share their political ideology, suggesting that aversion to uncertainty predisposes individuals to process political information in categorical, black-and-white terms that are shared by their political in-group (van Baar et al., 2021). In a follow-up study, de Bruin and colleagues (2023) found that interbrain synchrony while watching political debates correlated with similarity in neural patterns when participants were presented politically-relevant words (e.g., abortion, freedom, immigration) one at a time. The neural patterns in response to reading individual words is thought to reflect the semantic associations that are activated by the word. Thus, these results suggest that individual differences in the semantic representations of political concepts shape interbrain synchrony when viewing rich, real-world political content.

Collectively, this body of work demonstrates the varied ways in which interbrain synchrony is shaped by the interaction between the stimulus, individual predispositions and the social context.

Joint Attention

Fluctuations in interbrain synchrony has been shown to track fluctuations in attentional engagement. Song and colleagues (2021) examined the interbrain synchrony as participants watched a movie clip and listened to an audio story. A separate group of participants rated their level of attentional engagement as they watched the clip and listened to the story. In both cases, interbrain synchrony was stronger during moments that consistently were rated as highly engaging, suggesting that interbrain synchrony provides a marker of attentional engagement. Relatedly, other studies have found that emotional stimuli, which are known to capture attention, are more likely to synchronize neural activity (Nummenmaa et al., 2012; Trost et al., 2015).

A highly engaging stimulus can capture participants' attention, driving greater synchrony across brains. Attention can also be coordinated via social cues. The capacity to socially coordinate attention emerges as early as 6-months of age (D'Entremont et al., 1997), and is thought to be important for the development of language and theory-of-mind (Moore et al., 2014). In developmental studies, this is commonly studied by assessing infants' tendency to monitor and follow the target of others' gaze and gestures. In adults, observing that someone is looking in a particular direction automatically directs attention towards the corresponding location in space (Driver et al., 1999; Langton & Bruce, 1999). Gaze-following has been associated with interbrain synchrony between an observer and the target whose gaze they are tracking (Bilek et al., 2015; Dravida et al., 2020), suggesting that joint attention induces neural synchrony.

Interbrain synchrony has also been shown to be enhanced by direct eye-contact between two interacting partners (Hirsch et al., 2017; Kinreich et al., 2017; Leong et al., 2017). For example, interbrain synchrony is stronger between infants and adults if the adult is looking directly at the infant than if the adult is interacting with the infant with an averted gaze (Leong et al., 2017). Similarly, Hirsch and colleagues (2017) showed that interbrain synchrony was stronger during moments of eye-contact than when both interaction partners were directed to attend to the same stimulus. One interpretation of these results is that eye-contact serves as a "synchronization trigger" that resets ongoing neurocognitive processing to facilitate alignment with an interaction partner, though further work that experimentally manipulates eye-contact is needed to test the causal relationship between eye-contact and interbrain synchrony.

Communication

Interbrain synchrony is often studied in the context of communication (Jiang et al., 2012, 2015; Nguyen et al., 2022; Piazza et al., 2020; Schippers et al., 2010; Stephens et al., 2010). For instance, neural activity synchronizes between a speaker telling a story and participants listening to the speaker. Listeners whose neural activity were more strongly synchronized with the speaker performed better on a comprehension test on the story, suggesting that stronger interbrain synchrony was associated with successful communication (Stephens et al., 2010). Similar results have been observed in gestural communication – during a game of charades, activity in mentalizing areas of the guesser tracks activity in the motor cortex of the gesturer (Schippers et al., 2010). In these studies, listener/guesser (i.e. "receiver") brains are not perfectly aligned with that of the speaker/gesturer (i.e. "sender"), but instead lags several seconds behind, consistent with information flow from sender to receiver.

The popularization of fNIRS as a portable alternative to fMRI provides new opportunities to study interbrain synchrony during dynamic, real-time communication in groups. In one such study, Jiang et al. (2015) had participants engage in a discussion in groups of three. External raters then watched a video recording of the interaction and rated each participant on leadership qualities. All participants exhibited interbrain synchrony, but synchrony was stronger between participants rated as the leader of the discussion and the two participants rated as followers, than between the two followers. The magnitude of leader-follower synchrony correlated with ratings of the leader's communication skills, suggesting that a better communicator is more likely to induce interbrain synchrony with their communication partners.

Future Of Interbrain Synchrony Research

The past two decades have seen a growing number of experiments take advantage of new analytical methods to study interbrain synchrony. These studies are beginning to chart out the conditions under which interbrain synchrony emerges, as well as the functional role it performs in shaping shared perspectives, joint attention, and successful communication. New methods have been developed to extend the framework of interbrain synchrony beyond co-fluctuating activity time courses to include co-fluctuating patterns of multivoxel responses (Chang et al., 2021; Chen et al., 2017) and functional connectivity (Finn & Bandettini, 2021). It should be noted that, thus far, most evidence relating interbrain synchrony to behavioral outcomes has been correlational. An important direction for future research is to causally test the role of interbrain synchrony in facilitating social interactions. The incorporation of methods such as transcranial magnetic stimulation (Novembre & Iannetti, 2021) and real-time neurofeedback (Müller et al., 2021) are particularly promising in establishing a causal link between interbrain synchrony and behavior.

SOCIAL CONNECTIONS

Affiliation and supportive social relationships are vital for an individual's physical and psychological welfare across their life span. Human beings are intrinsically motivated to form and maintain lasting, positive, and significant relationships. Disruption of social connections and chronic social isolation are serious risk factors for morbidity and mortality, even after controlling for known risk factors, social status, and baseline measures of health (Holt-Lunstad, 2021). From birth until our last breath, we are dependent on others to survive, whether they are loved ones or strangers. This need for affiliation and positive evaluation from other conspecifics is found in all of us and can be considered a natural and universal phenomenon. We are also sensitive to prosocial acts and reciprocity in exchanges, and we derive many advantages from signaling to others that they are useful and contribute to the community's life.

Evolutionary Roots

In non-human primates, the work of biological anthropologists has demonstrated that positive social relationships contribute to the evolutionary success of many species, particularly in the baboons of the Okavango Delta in Botswana and the Rift Valley in Kenya (Cheney et al., 1986). These anthropologists used the word "friendship" to describe close, affiliative social relationships that require time, support, tolerance, loyalty, security, and reciprocity in social exchange. While

human and primate relationships are not exactly the same, it would seem that there are similarities that may reveal some essential aspects of the characteristics of friendship and may provide its ultimate explanations. Is friendship a by-product of kin selection, and a residual advantage? Or, is friendship the product of reciprocal altruism and favored by natural selection because it provides a basis for alliance formation? If natural selection has favored the ability to form friendships, then three aspects should hold: 1) friendship should increase reproductive success; 2) individuals should seek out friends (or try to); and 3) individuals should have the social intelligence to select the best partners. The results of many studies with primates are clear: having more and better "friends" is significantly related to reproductive success (Silk et al., 2010). In addition, quality social bonds matter more than social rank and status in surviving and thriving (Silk et al., 2003). Friendship has been observed not only in baboons, chimpanzees and other primates, but also in elephants, hyenas, whales and dolphins. Overall, there is growing evidence that natural selection favors the formation of stable positive social bonds associated with friendship relationships.

Biological Mechanisms Of The Links Between Sociality And Health

Like all social species, humans must find an ecological niche to survive, develop and ensure their well-being. The positive protective effect of one individual on another begins at birth with co-regulation between a mother and her offspring and is based on the reciprocal bonds provided by attachment. This requires not simply the presence of others but also the presence of significant others whom they can trust and with whom they can plan, interact, and work together to survive and thrive (Cacioppo & Patrick, 2008). This fundamental biological adaptation has a direct causal effect on health and well-being, regardless of stressful conditions. Throughout life, the presence and proximity of others helps to maintain costly somatic and metabolic resources (Beckes & Coan, 2011). Social networks in general provide reliable positive experiences and a stable and comforting set of social relationships. Such social connections also play a protective role in the face of physical, economic, and emotional challenges or injuries (Umberson & Karas Montez, 2010).

Impact Of Positive Relationships

We know from both common sense and empirical studies that satisfying interpersonal relationships have a robust positive association with happiness and health. These relationships even reduce mortality risk (Umberson & Karas Montez, 2010). Several explanations have been suggested to account for the link between positive relationships and health. One, in particular, focuses on the well-known link between social support and health. For example, partners in a quality relationship can help us to calm down when things are not going well. People in positive relationships may not argue as frequently or as intensely, but provide emotional and financial support to one another. Individuals who are engaged in positive and secured relationship gain a sense of identity, belonging, predictability and stability, and self-worth fostering a sense of meaning and purpose in life (Cohen, 2004).

This positive interpersonal process is a social dynamic in which one person's thoughts, feelings, and behavior alter those of another, and is fueled by positive emotions (Algoe, 2019). Emotions also promote social interaction by synchronizing brain activity across individuals, especially in the dorsal attention network, together facilitating interpersonal understanding (Nummenmaa et al., 2017). For example, shared laughter is a social phenomenon, as it occurs more often in the presence

of others and is highly contagious. It triggers activity in the pre-supplementary motor area (pre-SMA) where neurons when electrically stimulated, produce laughter sounds and smiles combined with feelings of happiness (Krolak-Salmon et al., 2006). Laughter also has been found to link to relationship satisfaction between romantic partners. In one study, couples in a relationship were asked to talk about their first meeting, which elicited a lot of positive emotions and laughter (Kurtz & Algoe, 2015). Their behavior – the actual time spent laughing, either together with their partner or alone– was videotaped. The extent to which couples shared laughs and laughed alone positively correlated with their feelings of closeness and safety towards each other.

Laughing has many positive physiological effects (Kljajevic, 2021). It decreases stress hormones such as epinephrine and cortisol and increases mesolimbic dopaminergic activity as well as pleasure and reward circuit activity. It also triggers endorphins, which are polypeptides synthetized by the pituitary and hypothalamus, and which are present in the brain and spinal cord. These hormones bind to morphine receptors and cause feelings of relaxation, well-being, and, in some cases, euphoria. Endorphins also help to manage stress and reduce pain sensations, and overall have an analgesic, anxiolytic, and relaxing action.

Several studies suggest that the mere presence of people we like and trust results in several co-occurring effects that cause a decreased physiological response to stress. Vigilance in the face of a threat while in the presence of people we like and trust is less necessary than when we are alone and, in keeping with the social baseline theory, it is less taxing on metabolic resources (Beckes & Coan, 2011). For instance, one fMRI study brain measured brain activity in participants who underwent a threat-of-shock paradigm under three conditions: 1) while alone; 2) while holding a stranger's hand, and 3) while holding a friend's hand. Handholding in general mitigates the neural threat response, particularly when the hand-holder is a familiar relationship partner (Coan et al., 2006). A significant decrease in hemodynamic activity was found in the neural circuits typically associated with emotion regulation (e.g., dorsolateral prefrontal cortex, or dlPFC) and the anterior cingulated cortex when holding hands with a partner relative to an alone condition. Responses to pain stimulation can be socially modulated as well. Across 26 studies, a systematic review documented that positive verbal and non-verbal interactions or positive interpersonally relevant primed interactions reduced pain sensitivity, whereas negative or ambiguous interactions led to increases in measures of pain intensity and unpleasantness (Krahé et al., 2013). Moreover, this meta-analysis found that the type of relationship between the participant and social partner moderates how interpersonal factors affect pain sensitivity with pain being more attenuated when the participant was receptive to support and knew that the partner was positively oriented toward them.

Finally, a large body of evidence demonstrates that social support is related to the autonomic nervous, neuroendocrine, and immune systems in ways that confer protection against disease (Uchino, Smith, Birmingham, & McKenzie, 2011). Overall, this reduces allostatic load, which reflects the wear and tear on the body due, in part, to chronically overworked physiological systems engaged in stress responses (McEwen, 2012).

The Lack Of Social Connection

A prolonged absence of social connection, as measured by perceived loneliness, has deleterious effects that are expressed at all levels of biological organization, from gene expression to psychologal well-being (Cacioppo & Hawkley, 2009).

In the 1970s, epidemiologists and sociologists began to establish a link between social relationships and health. A meta-analysis that include more than 3 million people indicated that social isolation and loneliness increase the risk of early mortality by 30% (Holt-Lunstad et al., 2015). Individuals who suffer from loneliness face health risks similar to those associated with obesity, diabetes, or smoking. Socially isolated people even feel less happy and die younger than their socially connected counterparts.

In species from insects to humans, a large body of empirical research demonstrates that the health of social animals is threatened when they are on the periphery of their society (Cacioppo & Hawkley, 2009). Social isolation reduces sleep, increases cellular stress, and decreases the lifespan and health of the fruit fly (Vora et al., 2022). It promotes obesity and type 2 diabetes in mice and humans (Nonogaki, Nozue, & Oka, 2007). Social isolation increases basal cortisol concentrations and decreases lymphocyte proliferation to mitogens in pigs (Kanitz et al., 2004). It modulates neuroinflammatory responses and exacerbates the size of an infarct or edema after a stroke (Karelina et al., 2009). It decreases the expression of genes regulating the glucocorticoid response in the frontal cortex of piglets (Poletto et al., 2006). In humans, loneliness and social isolation increase anxiety, and the activity of the hypothalamic-pituitary axis (the neuroendocrine system that controls stress responses and regulates many biological processes including digestion, the immune system, mood and emotion, sexuality, and energy storage and expenditure), and cause sleep fragmentation (Cacioppo & Patrick, 2008). People suffering from loneliness experience negative feelings, and notice unpleasant concepts more quickly and intensely than others, which tends to increase their anxiety and promote the onset of depression (Cacioppo, Norris, Decety, Monteleone, & Nusbaum, 2009).

Chronic stress, on the other hand, can have long-term physiological costs (McEwen et al., 2015). Chronic stress activates the autonomic nervous, neuroendocrine, and immune systems and increases the risk of cardiovascular disease. Chronic stimulation of the hypothalamic-pituitary-adrenal axis results in an imbalance of anabolic-catabolic hormones, which results in:

- Elevated morning and evening cortisol levels
- Elevated resting systolic blood pressure
- Increased risk of hypertension
- Excessive visceral fat storage
- Increased insulin resistance, leading to pre-diabetes
- Impact brain structure in specific regions including the amygdala and hippocampus

Allostatic load, which is "wear and tear on the body," is the secondary result of stressors (McEwen, 2012). This load accumulates when an individual is exposed to repeated or chronic stress.

Several functional neuroimaging and eye-tracking studies suggest that loneliness is related to an increase in selective attention in detecting social threats (Bangee et al., 2014 but see Saito et al., 2020 for contradictory results), combined with a less rewarding interest in positive social context via the reward circuit. In an fMRI study, participants who suffer from loneliness showed a reduced neural response in the ventral striatum (involved in the regulation of motivated behaviors and anticipation of rewards) when exposed to pleasant social images, and increased activation in the visual cortex directly related to their degree of perceived loneliness when exposed to negative social images (Cacioppo et al., 2009).

Additionally, loneliness has a discernable impact at the molecular level. A study conducted by a team of social neuroscientists and immunologists at UCLA found that the pattern of immune cells

called myeloid cells (a type of immune cell) is significantly different in people who are very lonely compared to those who are not (Cole et al., 2007). Lonely people have an unusually low number of myeloid cells, which makes them more likely to be vulnerable to viral infections. These individuals also carry an abundance of a second type of myeloid cell that stimulates the activity of genes that cause inflammation, and as a result, individuals who suffer from loneliness experience more inflammation than controls.

Such correlations are interesting. However, they do not explain any causal linkage. Can these findings be explained by loneliness or by the myeloid response? These same authors of this study repeatedly measured volunteers' subjective perceptions of social isolation while simultaneously monitoring their gene expression patterns and other physiological changes through blood samples. Results show that initially, the volunteers' feelings of isolation coincided with an increase in the activity of their inflammatory genes and a concomitant increase in the circulation of monocytes (i.e., immature immune cells) that are involved in inflammation and circulate in the brain and promote anxiety. Results also indicated an increase in signaling molecules associated with both inflammation and behaviors such as social withdrawal, feelings of suspicion toward the outside world, and a tendency to act more defensively in making low-risk decisions. This promotes new feelings of loneliness, which, in turn, triggers a new myeloid response. Thus, it seems that although loneliness can quickly take on a physiological path of its own, it can also create a state of chronic loneliness that is difficult to escape when things return to normal. As the myeloid feedback loop makes individuals more defensive and suspicious, the mere presence of others is not enough to restore the status quo - something else is needed.

In a second set of experiments, the same team of social neuroscientists encouraged healthy volunteers to perform simple acts of kindness to their fellow human beings, such as helping a colleague with an Internet issue (Nelson-Coffey et al., 2017). The researcher collected blood samples to examine myeloid responses. Interestingly, participants who were kind to others on a weekly basis had precisely the opposite gene expression of that previously observed in people suffering from loneliness. Encouraging people who suffer from loneliness to engage in acts of kindness towards others, including via the internet, would reduce both the feeling of loneliness and gene expression at the molecular level.

Overall, the work in animals and humans demonstrates that the association between social isolation and health is not simply the result of behavioral changes. Perceived social isolation and the subsequent modulation of neural, hormonal, cellular, and genetic processes play a role in impaired health and well-being (Norman et al., 2012). It is possible to change this vicious circle by thinking of others and being generous with one's time, smile or money. The important thing, and probably the most difficult one, is to be the agent that spontaneously and voluntarily initiates such actions.

SOCIAL NETWORKS

People’s lives are embedded within intricate webs of social relationships and interactions. These relationships and interactions have a powerful influence over nearly every aspect of daily life, including health, well-being, and professional success. The study of social networks is an endeavor that spans sociology, psychology, and computer science, as well as the interdisciplinary field of network science, which specializes in the study of complex networks across physical, biological, and sociological domains (Borgatti et al., 2009). Social neuroscientists combine tools from network

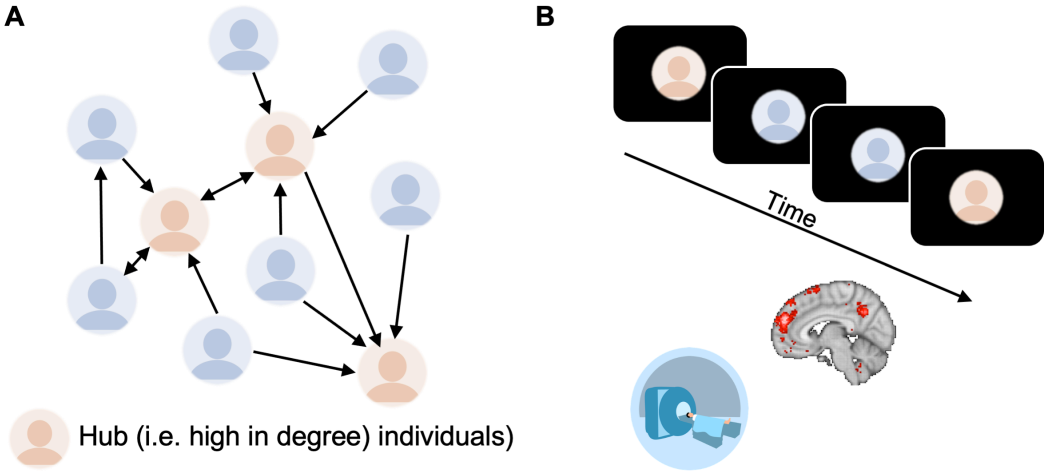
science with neuroscience methods to examine the neural and mental architecture of how people represent and navigate social relationships (Baek et al., 2021). This multimethod approach holds great promise in advancing our understanding of how people shape, and are, in turn, shaped by social ties.

A Network Science Approach To Studying Social Networks

Networks of social relationships can be represented as a "network graph" where each individual is a "node" that is connected to other individuals by "edges". Network graphs can be mapped out by having individuals nominate other individuals based on a metric of interest (e.g., friendship, popularity, attractiveness; Morelli et al., 2017) or by collecting objective measures of connection strength (e.g., number of co-authored articles between scientists; Glänzel & Schubert, 2005). Depending on the research question and the type of data used to construct the network graph, a network can be "directed" or "undirected." A directed graph is a graph where the edges have directionality. For example, in a network of popularity nominations, a directed edge from person A to person B indicates that person A nominated person B as popular but does not imply that person B also nominated person A as popular. In contrast, an undirected graph is a one where the edges do not have directionality. In a network of article co-authorship, for instance, the edges are based on a shared attribute between two individuals, and thus have no directionality.

Various mathematical and statistical tools have been developed to analyze and visualize both directed and undirected graphs. These methods are agnostic to the type of data used construct the network, and have been applied to physical networks (e.g., how subway lines are connected in a city), biological networks (e.g., how brain regions interact with one another), and semantic networks (e.g., how the meaning of words relate to one another). By transforming social relationships into a network graph representation, social neuroscientists can capitalize on these methods to compute useful *sociometric* measures about each individual in the network or the network as a whole. These measures can then be used as variables of interest in neuroimaging analyses to identify regions of the brain that encode information about social relationships.

***Figure 2*: How Social Networks Are Represented in the Brain**



Note: Example experiment investigating how social networks are represented in the brain. A. The experimenter first maps out the social network of a community by having every individual in the community nominate members of a community on a particular trait (e.g., who are your

closest friends?). These nominations are used to construct a graph, which can be used to calculate sociometric measures about each individual in the network. For example, we can identify hub individuals who are frequently nominated by their peers. B. A subset of the participants then have their brain activity measured while viewing photos of every individual in the network. The experimenter can then examine how brain activity changes as a function of the sociometric property of the viewed individual.

The Neural Representation Of Social Networks

In an early study that combined social network analysis and functional neuroimaging, Zerubavel and colleagues (2015) had members of two student organizations rate how much they liked each of their fellow group members. The ratings were then used to generate a network graph where each edge represented how much an individual was liked by a particular group member. This allowed the researchers to obtain a measure of the popularity of each group member by computing that strength and number of edges directed towards that individual (i.e. the individual's "indegree centrality"). A subset of participants was then recruited into an in-person experiment, where they viewed photographs of their fellow group members while undergoing fMRI. Activity in regions associated with affective valuation, such as the ventromedial prefrontal cortex (VMPFC), ventral striatum and amygdala, as well as regions associated with theory-of-mind such as the dorsomedial prefrontal cortex, precuneus, and temporoparietal junction correlated with the popularity of the viewed target. Importantly, these analyses controlled for how much the participant personally liked the target. In other words, activity in these regions tracked how much the group liked the target collectively, above and beyond the participant's personal liking of the target.

Popularity is only one of the many sociometric properties represented in the brain. Using a similar approach, Morelli et al. (2018) showed that fine-grained patterns in the same affective valuation and theory-of-mind regions while viewing photographs of dorm mates could be used to predict the extent to which a target individual was perceived as socially valuable by other members of the dormitory. In a different study, Parkinson et al. (2017) assessed if different sociometric properties were independently encoded in the brain. Similar to the other experiments, participants viewed photographs of others in their community while undergoing fMRI. For each target, the researchers computed the "degrees of separation" (social distance) to the participant in the scanner, how well the target was connected to well-connected others (eigenvector centrality), and the extent to which the target serves as a bridge between otherwise unconnected individuals (brokerage). Each of these properties were then entered as variables in a linear regression model to assess the extent to which they explain variations in brain activity while participants viewed photographs of the target. The analysis indicated that social distance and brokerage were encoded in different parts of the superior temporal cortex, while eigenvector centrality was encoded in an extensive network of brain regions that included the medial prefrontal cortex, precuneus and temporoparietal junction.

Collectively, these studies indicate that people encode sociometric information about their peers. This information is encoded most prominently in regions associated with affective valuation and regions associated with theory-of-mind. Notably, participants in these studies were not instructed to think about characteristics of the target while viewing the photographs. Instead, they performed a visual detection task unrelated to social cognition (e.g., respond when a photograph repeats itself). Thus, these results suggest that sociometric information is automatically retrieved when encountering a familiar other, which may in turn guide subsequent social inferences and behavior. The functional role of the spontaneous retrieval of sociometric information during person

perception remains an underexplored question. Future work that studies social decision-making in social networks would help shed light on how the activation of sociometric information influences social interactions. Jiang and colleagues (2023) conducted one such study where participants performed a decision-making task while embedded within an experimenter-created social network. In addition to performing the task themselves, participants had the opportunity to observe the decisions of their immediate neighbors in the social network. The researchers found that participants learned from the decisions of others, and were more strongly influenced by neighbors who were more well-connected. Activity in the dorsal anterior cingulate cortex (dACC) correlated with the connectedness of a neighbor and predicted the degree to which participants' choice was influenced by the neighbor. These results suggest a neurocomputational mechanism by which individuals weight information they receive according to the social network position of the source of that information.

Learning The Structure Of Social Relationships

To learn the public transport network of a city, a traveler can study the map provided by the city. People, however, do not have that luxury when trying to learn the complex network of social relationships in their lives. There exists no map or guidebook that lays out one's social network in clear and simple terms. Instead, people have to piece together the structure of their social network via small, often dyadic, interactions with the people around them. Nevertheless, people can exploit structure in their social environment to learn and represent social relationships with seeming ease (Lau et al., 2018; Son et al., 2021). How does the brain support this impressive feat?

Studies highlight the role of the hippocampus and the adjacent entorhinal cortex in learning a "cognitive map" of social relationships (Schafer & Schiller, 2018). Both the hippocampus and entorhinal cortex are brain regions in the medial temporal lobe of the brain, and have long featured in studies of spatial learning (Moser et al., 2008). In particular, "place cells" in the hippocampus signal an animal's specific spatial location in space, while "grid cells" in the entorhinal cortex organize space in a grid-like fashion. Together, place cells and grid cells encode an internal coordinate system (i.e. a "map") of where an individual is currently located in physical space relative to other spatial landmarks. Several studies suggest that the brain may rely on the same neural systems involved in navigating physical space to navigate *social space* (Park et al., 2021; Schafer et al., 2022; Tavares et al., 2015; Zhang et al., 2022).

Given the difficultly of tracking the formation of real-world social relationships over time, most studies have utilized experimental paradigms where people interact with fictitious individuals in laboratory-constructed social networks. This approach allows researchers the opportunity to examine how people learn the structure of social relationships in a controlled and accelerated setting. For example, Tavares et al. (2015) had participants play a "choose-your-own-adventure"-style role-playing game while undergoing fMRI. Participants played the role of an individual who had relocated to a new city and had to find a new home and job through virtual interactions with fictitious characters. Through the decisions made by the participant, these characters can be organized in a 2-dimensional space of relative power (dominant-submission) and affiliation (friendly-hostile). This work draws directly from theories in social psychology proposing that people describe other people along these two dimensions (Wiggins et al., 1989). The dimensions can also be mapped, albeit not perfectly, to other two-factor models of person perception in social psychology, including the competence and warmth dimensions of the stereotype content model

(Cuddy et al., 2008) and the dominance and trust dimensions of the face evaluation model (Todorov et al., 2008).

Activity in the hippocampus dynamically tracked the coordinates of the characters within the two-dimensional space as participants progressed through the narrative, indicating that the hippocampus was sensitive to the relative power and affiliation of the characters in the fictitious narrative. As the characters change in relative power and affiliation based on the narrative, so does the response of the hippocampus to the characters (Tavares et al., 2015). A follow-up study found that patterns of activity in the hippocampus were more similar when interacting with individuals who were more similar in relative power and affiliation, providing further evidence that the hippocampus encodes information about social relationships (Schafer et al., 2022). Specifically, these results suggest that just as the hippocampus encodes a specific location in *physical space* during spatial navigation, it may also encode a specific location in *social space* during social navigation.

Conceptualizing social relationships within the framework of cognitive maps allows social neuroscientists to draw on a broader literature from which to formulate new hypotheses and draw on new methodological approaches to study how people navigate social relationships. For example, a grid-like representation of physical space is thought to facilitate the brain's ability to make efficient spatial inferences (Bellmund et al., 2018). To test if social space is also organized in a grid-like fashion, Park and colleagues (2021) had participants learn about a set of entrepreneurs who vary along competence and popularity. Participants then performed a decision-making task involving the entrepreneurs while undergoing fMRI. Neural responses in the entorhinal cortex while participants performed the task revealed a two-dimensional grid-like representation of the entrepreneurs. The researchers also found similar grid-like coding of social space in medial prefrontal cortex, superior temporal sulcus, anterior temporal lobe, and temporoparietal junction. These additional regions have been widely associated with social cognition, and the grid-like representation may reflect modulation from the entorhinal cortex when navigating social knowledge. Future work that directly compares spatial and social navigation would allow researchers to tease apart whether the involvement of these regions is, indeed, specific to social contexts.

Altogether, these studies suggest that the brain relies on similar neural processes to navigate physical and social spaces. Thus, the use of spatial descriptions to describe social relationships (e.g., "her position is higher than mine," or "he is close to me") may be more than just an analogy, and could instead be revealing fundamental truths about how the brain encodes social relationships.

SOCIAL HIERARCHIES

Social hierarchies are ubiquitous across social species and have a profound impact on how individuals perceive and interact with others. Social hierarchies are present early in the children's development across cultures (Hawley, 2015) and are thus not strictly a product of top-down cognitive control, but rather emerge naturally. At 17 months of age, infants expect a leader who observed a within-group transgression to intervene and rectify it but would not have such an expectation from a non-leader who observed the same transgression (Stavans & Baillargeon, 2019). Social status, the relative rank of an individual along one or more dimensions within a given social hierarchy, is both dynamic and derived from multiple sources and social dimensions. The relative importance of these dimensions depends on specific characteristics of the individual and the social context. Neural circuits engaged in the perception of social status piggyback on systems that

evolved for selective attention, person perception, computing rank order comparison, and valuation. Change or rank in social status affects the immune system, gene expression, the neuroendocrine system as well as activity in specific brain circuits.

The groups in which social animals live have a complex internal structure. Their size and composition vary depending on ecological pressures and niche construction, but they have one characteristic in common: a social hierarchy. Indeed, all social species are organized into hierarchies in which individuals within have asymmetric access to resources. Social species show a natural tendency to organize themselves into a hierarchy where higher-ranking members have more power, influence, and access to resources and mating opportunities. In non-human animals, those who occupy the highest positions are called "dominants" or alpha females/males. In humans, they are referred to in many ways with examples such as elites, the upper or superior class, the aristocracy, celebrities, and influencers. Contrary to some other primate species, the rank in the social hierarchy of humans is not inherited but is a provisional outcome of dynamical processes involving social learning and competitive interactions within social contexts.

Human hierarchies are not organized in a linear, one-dimensional way. Rather than being determined and maintained by aggressive and affiliative physical encounters, hierarchies among humans are established along various dimensions based on skill, reputation, economic, and intellectual status. However, human status hierarchies are not only based on dominance and agonistic behavior, but also on consensual deference to legitimate leaders and competent experts (Thomsen, 2020). The structure of human hierarchies is thus highly flexible, multidimensional, and dependent on multiple contexts, history, and other related factors.

However, previous evolutionary adaptations are always present. In most species, there is a direct relationship between social status and reproductive success in that higher-status individuals are less likely to die from predation and starvation. These individuals have better access to resources, and mating partners, and are more likely to leave many offspring, showing a direct link between social status and inclusive fitness. Social hierarchies dynamically emerge because they are an effective way to maximize group cohesion, minimize aggression, define social roles, promote appropriate interactions, facilitate social learning and allocate scarce resources. As a cultural species, learning from each other plays an adaptive role, and explains why we automatically and unconsciously attend from other conspecifics based on cues of prestige, success, and skills (Henrich, 2016).

People are sensitive to a suite of signals or behaviors that serve to establish and convey social status (Koski et al., 2015). Some of these signals seem to be conserved between nonhuman primates and humans, especially nonverbal cues such as body size and posture, gaze, and voice. People in high social positions signal status, power and wealth via multiple channels, and such signals have an effect on others' cognition and behavior -even the temporary experience of social status influences behavior. As an illustration, one study required female undergraduates to walk around campus while carrying a luxury or a non-luxury purse (Wang et al., 2015). Afterward, they were given an opportunity to act more or less selfishly by taking the last pen from a common pen box. Participants who had been temporarily imbued with a higher social standing (by carrying the luxury purse in this experiment) were significantly more likely to behave selfishly. Relatedly, a different study showed that people are more likely to comply with a request (e.g., complete a short survey) when the experimenter is wearing a luxury brand sweater (Nelissen & Meijers, 2011).

Proximal Mechanisms Of Social Rank Perception

Perceiving, monitoring, and adapting to signals of social dominance is associated with biological responses at many levels. Endocrine responses play an important role in the development and expression of social status. In many species (including rainbow trout, zebrafish, and primates) there is a negative association between cortisol hormone levels and social rank, but high levels have been detected in dominant individuals, revealing a U-shaped association between stress hormone and social rank (Dwortz et al., 2022). Serotonin levels also increase in subordinates who improve their social status (Watanabe & Yamamoto, 2015). Serotonin is involved in the establishment and maintenance of status with higher individuals being associated with affiliative behavior and good social skills and is correlated with androgen and serotonin levels (Edwards & Kravitz, 1997).

Interestingly, periods of instability that result in changes in the social hierarchy cause a spike in cortisol levels and marked modifications of other hormones levels in all members of a social group. For instance, following competition, losing individuals show a decrease in androgen levels (Oliveira et al., 2009). After a game, male players who won show elevated levels of testosterone compared to those who lost (Booth et al.,1989). Most interestingly, this is true even when the competition involves no muscle action, as in chess competition, with a rise in testosterone before the game that either stays elevated or decreased depending on the outcome (Mazur, Booth, & Dabbs, 1992). Administration of testosterone versus placebo in female participants increases utilitarian response to moral dilemmas and has a profound effect on neural activity. It reduces activity in the temporoparietal region and decreases functional coupling between the dorsolateral prefrontal cortex and amygdala (Chen et al., 2016). From these studies, it seems clear that testosterone plays an important role in maintaining and achieving a high social status.

Vicarious fluctuations in hormones levels have also been reported in sport fans. Testosterone levels in college basketball and soccer fans watching a game reflect those of the team they support (Bernhardt et al., 1998; Van der Mej et al., 2012). Testosterone levels increased among fans of winning teams and decreased among fans of losing teams. This effect is usually interpreted as supporting the challenge hypothesis, as testosterone levels of watchers increased to prepare their organism to defend or enhance their social status and self-esteem.

Brain systems supporting the representation of dominance hierarchy are highly distributed and involve a coordinated activity across different mechanisms involved in selective attention, person perception, computing rank order comparison, social memory, and valuation based on numerous sensory cues (visual, auditory, and more abstract ones). The informational output of this system guide individuals' decisions on how to interact depending on other conspecifics' social identity, as well as when to become more vigilant and aware of others' intentions and emotions.

A key aspect of navigating social hierarchies is knowing where to orient attention, with dominant individuals receiving disproportionate attention compared to subordinates. This requires computations of social ranking comparison in the intraparietal sulcus (IPS), which plays a crucial role in numerical magnitude judgment. The IPS also acts with the frontal eye fields to exert top-down influences on visual areas during the spatial orienting of attention. Functional MRI studies have consistently demonstrated that neurodynamic activity in the IPS is modulated by the implicit perception of social status signals inferred from objects, symbols (e.g., military medals), faces, or dominant bodily postures (Chiao et al., 2010 for review). In keeping with numerical comparison (Dehaene et al., 2003), a greater degree of activity is detected in that region when the hierarchical difference between two stimuli was closer than when the difference was greater.

Social information is represented in other functional systems as well. The medial prefrontal cortex, temporoparietal region (involved in reasoning about intentions, desires and beliefs), and the systems that code value and saliency (amygdala, striatum, and orbitofrontal cortex) are preferentially engaged when one perceives and interacts with people of high social status (Mattan et al., 2017 for a review). For instance, one study documented dissociable neural responses to perceived social rank in an interactive social context (Zink et al., 2008). In both stable and unstable social hierarchies, viewing a person from a higher rank was associated with differential involvement of the perceptual-attentional system and the striatum. In the unstable hierarchy condition, additional regions were recruited related to emotional saliency (amygdala), social cognition (medial prefrontal cortex), and behavioral readiness. The authors also reported that the striatal activity was greater when participants were informed of their win or loss and when they defeated a superior human player (the "social context" in this experiment). However, this activation did not occur when participants defeated a superior computer player (the "non-social context"). Finally, social hierarchical consequences of performance were neutrally dissociable and of comparable salience to monetary reward, providing a neural basis for the high motivational value of status.

Another study found that when participants are presented with faces previously associated with varied social status, selective activation in the orbitofrontal cortex is detected (Koski et al., 2017). The involvement of this value-sensitive region, which belongs to the extended face-processing network, seems to discriminate high- from low-status individuals. Importantly, this study did not prompt participants to attend to any status-relevant information, and therefore, the results are unlikely to be due to increased attention toward those signals. Similarly, electrophysiological recordings of ventral striatal neurons in rhesus macaques indicate that macaques assign more value when they perceive faces of high-status conspecifics and attend more to cues about their own mental states, which influences their motivation to act (Klein, & Platt, 2013).

Overall, several specific neural circuits are involved in the perception and formation of dominance and social hierarchy cues, and both are dependent on the individual's prior experience in the social context in which they occur.

Impact Of Socioeconomic Status On Health

Social hierarchies play a functional role in societies, but they can also have deleterious influences on individuals who rank low within the hierarchy. As an example, socioeconomic status (SES) is a very significant area of disparity in human societies. Previous research has shown public health to be strongly linked to social and economic circumstances (Wilkinson, & Pickett, 2006). Inequalities can have major psychological effects and many health problems are the result of increased social stress. Thus, often, the lower the degree of inequality, the healthier the society. A 2014 report from the World Bank concluded that in the developed world, it is not the richest countries that have the best levels of health and life expectancy, but the most egalitarian ones (as measured by the gap between the rich and the poor). For instance, Japan, Hong Kong, and Singapore scored the highest whereas the United States scored the lowest in economic equality.

The social environment has a powerful influence on the body and the brain through the neuroendocrine, autonomic, and immune system (McEwen, 2012), and the biological embedding of early experiences has long-lasting effects on cognition and behavior. A review of the empirical evidence suggests that childhood SES affects not only the outcome but also the pace of brain

development (Shonkoff, 2003; Tooley et al., 2021). Specifically, social experiences that are both negative and chronic are the most likely to accelerate brain development and reduce neural plasticity. Repeated exposure to adverse experiences and stress also leads to early maturation of the networks that process these experiences and augments glucocorticoid levels, allostatic load and inflammatory processes that age the entire body (McEwen, 2012). Low SES is accompanied with a host of anatomical and functional deficiencies including smaller hippocampal volumes, reduction of prefrontal gray matter, and greater amygdala reactivity of threatening stimuli (Hackman et al., 2010). Chronic activation of the stress response not only increases the risk of numerous diseases over the lifespan. It also leads to epigenetic transgenerational modifications of neurophysiological reactivity and maladaptive behaviors such as poor self-regulation and impulsivity.

Research with non-human primates shows that dominance rank influences the extent to which an individual tolerates physical and psychosocial stressors. Although humans are not hierarchical in a unidimensional way, as other primates are, findings from animal models still largely apply. In particular, SES gradient is a robust example of social inequality predicting patterns of disease (Sapolsky, 2005). Progressive descent in SES predicts increased risk of cardiovascular, respiratory, obesity, and psychiatric diseases, low birth weight, infant mortality, and more general all-cause mortality. This relationship is primarily due to the influence of SES on health, rather than the reverse (Wilkinson & Pickett, 2018). Furthermore, it has been demonstrated that subjective SES (the person's perception of their rank) provides a better predictive utility than objective measures (Cundiff & Matthews, 2017). A large study that included 23,931 individuals documented that SES differences are reflected in neuroanatomy at both cortical and subcortical levels (Kweon et al. 2022).

Overall, social hierarchies are ubiquitous in all social species. The organization of groups into hierarchies serves adaptive functions that benefit all members, and natural selection has shaped our perceptual, cognitive, hormonal, and neurophysiological systems to detect and flexibly adapt to social status signals. However, rank in hierarchies can also dramatically influence the quality of life and health.

MOTIVATED COGNITION

People tend to interpret information in a manner that supports desired conclusions, a phenomenon known as motivated cognition. For example, voters in an election might explain away the flaws of their preferred candidate. Similarly, scientists can fall prey to confirmation bias when drawing conclusions from data. Studies have begun to characterize the neural mechanisms underlying how desires, wants, and beliefs affect human reasoning and evaluative processes, including how motivation biases human cognition as early as visual perception, how the brain learns more positive outcomes than from negative outcomes, and how people form unrealistically optimistic expectations about future events.

What Is Motivated Cognition?

The term "motivated cognition" takes on different meanings across sub-disciplines within psychology. In behavioral and cognitive neuroscience, motivated cognition is a broad umbrella term that encompasses the effects of reward signals, emotion, and motivational states on cognitive processing (Braver et al., 2014; Gottlieb et al., 2016; Madan, 2017; O'Reilly, 2020). For instance,

animals and people performing better on a task when incentivized by reward has been described by behavioral and cognitive neuroscientists as a form of motivated cognition. In social and personality psychology, motivated cognition tends to be defined in narrower terms, referring to a consistent bias in cognitive processing that steers people towards desired conclusions (Dunning, 2015; Higgins & Molden, 2003; Hughes & Zaki, 2015). Despite calls to broaden the scope of motivated cognition to include *non-directional* motivational effects (Kruglanski, 1999), social and personality psychologists have generally used the term to refer to a *directional bias* towards desired conclusions.

Inconsistency in how the term is used can be confusing when researchers trained in one tradition encounter literature from a tradition with which that they are less familiar. Given the inherently interdisciplinary nature of social neuroscience, this is particularly problematic for collaborative work across cognitive neuroscience and social psychology. As such, we find it important to clarify the scope of motivated cognition as we begin this section. Specifically, we adopt the narrower definition of motivated cognition – using the term to refer to a directional bias in cognitive processes to arrive at desired conclusions. Motivated cognition, according to this definition, includes phenomena such as wishful seeing, self-enhancement effects, unrealistic optimism, and motivated reasoning. The combination of tools and perspectives from social psychology and cognitive neuroscience have led to new insights on the nature of these biases, as well as how they are implemented in the brain.

Wishful Seeing

Wishful seeing is a specific case of motivated cognition where people are biased towards seeing the world in a way that fits with their goals, desires, and needs (Dunning & Balcetis, 2013). For example, sport fans might be predisposed towards seeing an ambiguous foul on the field in a manner that favors the team they support. In the laboratory, participants presented with ambiguous visual images with multiple potential interpretations (e.g., a line drawing that could be seen as either a horse or seal) tend to report seeing the interpretation associated with a desirable outcome (Balcetis & Dunning, 2006; Leong et al., 2021). Other studies have found that desired objects are judged as closer (Balcetis & Dunning, 2010), and a glass of water is reported as larger when an individual is thirsty (Veltkamp et al., 2008). These results are often interpreted as an influence of one's internal states on perceptual processes, biasing people to see the world not as it is, but how they wish it to be.

The notion that internal states shape perception was popularized by the "New Look" movement in psychology during the 1950s. The claim, however, has long been the subject of methodological critiques and controversy (Erdelyi, 1974; Firestone & Scholl, 2016). A common criticism is that wishful seeing might reflect biases in subjective reports rather than subjective experience. That is, goals and desires could be biasing what people report seeing (i.e. a bias in judgment), rather than what they see (i.e. a bias in perception). Without direct access to people's perceptual experience, it has been difficult to demonstrate conclusively motivational effects on visual perception. As such, the debate on the nature of these effects has continued to the present day.

Voss and colleagues (2008) proposed the "perceptual" and "judgmental" components of a bias can be dissociated using a drift diffusion model (DDM). A DDM is a computational model commonly applied to study the cognitive processes underlying binary decisions in mathematical terms (Ratcliff & McKoon, 2008). In a DDM, a perceptual judgment is modeled as the accumulation of

sensory evidence towards one of two decision thresholds. The starting point and rate of sensory evidence accumulation are free parameters that can be estimated by fitting a DDM to choice and response time distributions to experimental data. The DDM has been shown to account for both human behavior and neural activity across a wide range of paradigms (Forstmann et al., 2016).

Within the DDM framework, a systematic bias in perceptual judgment might reflect a shift in the starting point of evidence towards the threshold of one of the decisions. This reduces the amount of evidence needed to make the corresponding decision without affecting the accumulation of sensory evidence, and is consistent with a "judgmental bias" independent of sensory processing. A bias can also affect the accumulation of sensory evidence by weighting evidence for one percept more than the other. Biases in the rate of sensory evidence accumulation would thus be consistent with a "perceptual bias".

To test if wishful seeing arose from a judgmental bias or a perceptual bias, Voss and colleagues (2008) fit a drift diffusion model to data from a color categorization task where participants were rewarded for correctly categorizing whether an image contained more orange or green pixels. In addition, participants would earn more money if the image was of a particular category, and were thus motivated to see that category over the other. Crucially, the additional "category bonus" was not contingent on participants' judgments – they would win the additional money as long as the image objectively belonged to the incentivized category, irrespective of how they categorized the image. As such, participants would earn the most money if they ignored the category bonus and categorized the images accurately. Nevertheless, the authors found that participants were more likely to categorize the image as the color category they were motivated to see, and that this bias was due to *both* a bias in the starting point and a bias in the rate of sensory evidence accumulation. These findings suggest that wishful seeing reflects a bias in judgment as well as a bias in perception.

Functional neuroimaging has been used to provide convergent neural evidence for these modeling results (Leong et al., 2019; Calabro et al.,2022). In an fMRI study, Leong and colleagues (2019) presented participants with composite images created by averaging a face image and a scene image in different proportions, and rewarded participants for correctly categorizing whether the face or scene image was of greater intensity. As before, participants were financially motivated to see one category over another. Prior work in the cognitive neuroscience of object perception had identified distinct patterns of neural activity associated with the visual processing of faces and scenes (Grill-Spector, 2003). Thus, the authors were able to obtain a neural measure of perception by quantifying the level of face and scene neural activity in the brain. They found that wanting to see a category enhanced the neural representation of that category, suggesting the motivation to see the category shaped how the visual system encoded the presented image, consistent with a bias in perception. The magnitude of the motivational enhancement of category-specific neural activity tracked the rate of evidence accumulation estimated by a DDM, providing convergent validity of the model's ability to isolate perceptual biases. On the other hand, biases in the starting point were associated with activity in the nucleus accumbens, an area in the ventral striatum involved in shaping motor responses based on reward contingencies (Floresco, 2015).

The use of computational model and neuroimaging provides converging evidence that wishful seeing reflects *both* a bias in judgment and a bias in perception. The results also allow for a reframing of classic debates on the nature of wishful seeing. Instead of asking *whether* goals, desires, and needs bias perception, researchers can use these new tools to characterize *how* motivational factors affects the underlying cognitive and neural processes, thus providing a more mechanistic understanding of wishful seeing.

Motivated Self And Social Perception

Seventy percent of high school seniors rated their leadership skills as above average (Dunning et al., 2004), 80% of college students rated themselves as above average drivers (McCormick et al., 1986), and 90% of faculty members rated themselves as above average teachers (Cross, 1977). Across a wide range of desirable abilities, attributes and personality traits, people tend to think of themselves as better than their average peer (Zell et al., 2020). This phenomenon has gone by many names in the social psychology literature, including the *better-than-average-effect*, *illusory superiority*, *self-enhancement bias*, and the *Lake Wobegon effect*, after the fictional town where "all the women are strong, all the men are good-looking, and all the children are above average". Yet, for any given quality, only half of the population can truly be above average. In other words, there are many people who are convinced that they are better than they really are.

One mechanism by which people arrive at positive self-perceptions is via the active reinterpretation of negative information about themselves. For example, when acting in a selfish manner, one might justify their actions in a way that preserves their perception of themselves as kind and generous. Alternatively, positive self-perceptions can also arise from greater accessibility of positive information about the self. The distinction between these two hypotheses is that the former assumes positive self-perceptions are the result of the *controlled* processing, while the latter assumes that they are *automatic* and require little mental effort. Functional neuroimaging can help adjudicate between these hypotheses. If positive self-perceptions arise from controlled processes, we would expect an *increase* in activity in brain regions associated with cognitive control when people make positive judgments of themselves. In contrast, if positive self-perceptions are automatic, we would expect a *decrease* in activity in these regions. Thus, examining neural activity while participants form positive self-perceptions might allow researchers to distinguish between these two hypotheses.

To that end, Beer and Hughes (2010) had participants rate themselves on 200 traits while undergoing fMRI. Reproducing results from past behavioral studies, participants rated themselves as above average on positive traits and below average on negative traits. Activity in the orbitofrontal cortex and dorsal anterior cingulate cortex, brain areas associated with cognitive control (Ochsner & Gross, 2005), were *negatively* correlated with positive self-evaluations, suggesting positive self-perceptions are more likely when the brain is less engaged in controlled processing. Consistent with this interpretation, reducing the amount of time participants had to engage in controlled processing (by requiring them to respond more quickly) *increases* the magnitude of positive self-evaluations. A follow-up study found that people were also more likely to rate romantic partners as above average, suggesting that the positivity bias extends to close others. The positive evaluations of romantic partners were similarly negatively correlated with activity in the orbitofrontal cortex and the dorsal anterior cingulate cortex (Hughes & Beer, 2012). Altogether, these results suggest that above-average judgments of self and close others arise from automatic processes.

These findings, however, do not preclude a role of control processes in maintaining positive self-perceptions. In particular, control processes might be particularly important when confronted with negative evaluative feedback. In cases like these, control processes might be engaged to actively reappraise the negative information. In line with this hypothesis, individual with higher self-esteem exhibit increased activity in control regions when processing negative social evaluative feedback, suggesting that they might be more prone to reappraising the negative information

(Somerville et al., 2010; Schie et al., 2018). Thus, the cognitive and neural mechanisms that give rise to overly positive self-perceptions likely differ depending on the social context.

Unrealistic Optimism

People have a tendency to be unrealistically optimistic about the future – they often overestimate the likelihood of positive events and underestimate the likelihood of negative ones (Shepperd et al., 2013; Weinstein, 1980). Relative to their peers, people think they are *more* likely to do well in their career, live past the age of 80 or be in a happy relationship, and *less* likely to be fired from their job, develop serious illnesses, or experience a divorce. These optimistic beliefs engender positive feelings about the future, buffering one from anxiety and stress. However, they can also lead to problematic behavior. In particular, believing that a negative event is unlikely to happen (e.g., contracting COVID-19) reduces the likelihood that one would engage in preventive measures (e.g., wearing a mask) (Salgado & Berntsen, 2021).

Normative models of learning assume that people update expectations in the face of new evidence (Jacobs & Kruschke, 2011). If people have unrealistically optimistic beliefs about the future, they ought to update those beliefs when confronted with information indicating that their beliefs are overly positive. Yet, optimistic beliefs often persist in the face of contradictory information (Weinstein & Klein, 1995). To investigate the biological basis of the persistence of unrealistically optimistic beliefs, Sharot et al. (2011) had participants perform a belief-updating task while undergoing fMRI. Participants were asked to estimate the likelihood of experiencing a series of adverse life event (e.g., developing cancer, experiencing a robbery), and were then told the average probability that each event would occur to an average person in the same socio-cultural environment. Participants then provided a revised estimate of how likely the adverse event would occur to them. The authors found that participants updated their beliefs more when information was better than expected (e.g., the participant's estimate of developing cancer was higher than the average likelihood of developing cancer) than when information was worse than expected (e.g., the participant's estimate of developing cancer was lower than the average likelihood of developing cancer). In other words, participants' learning was *asymmetric*, with favorable information weighted more heavily than unfavorable information.

For each adverse event, the authors computed an "estimation error" as the difference between a participant's estimate and the actual likelihood of the event. When new information was better than expected, that is, when the participant's self-estimate of the negative event was higher than the actual likelihood, the estimation error would be positive. When information was worse than expected, the estimation error would be negative. An *optimal* learning agent aimed at maximizing belief accuracy would learn equally from positive and negative estimation errors. Participants' final estimates, however, were more consistent with an *optimistic* learning agent that learned more positive estimation errors than negative estimation errors. In the brain, this asymmetric learning was associated with the failure to encode negative estimation errors in the right inferior frontal gyrus (IFG) (Sharot et al., 2011), an area known to track expectancy violations (Sherman et al., 2016). These results suggest that unfavorable information is encoded less strongly in the brain, giving rise to optimistic belief updating. Interestingly, causally perturbing the left IFG reduced optimistic belief updating, but perturbing the right IFG had no effect (Sharot, Kanai, et al., 2012). Furthermore, structural connectivity between subcortical structures and the left IFG, but not the right IFG, (Moutsiana et al., 2015) was associated with optimistic belief updating. Further work will be needed to tease apart the distinct roles of the left and right IFG in optimistic belief updating.

At the neurochemical level, unrealistic optimism has been linked to the neuromodulator dopamine. Pharmacologically increasing dopamine levels in people has been shown to temporarily increase optimistic belief updating (Sharot, Guitart-Masip, et al., 2012). These results fit well with dominant theories of the role in dopamine in learning and decision-making. According to these theories, dopamine neurons encode a "reward prediction error" signal important for driving the updating of reward expectations (Glimcher, 2011; Niv, 2009). When outcomes are better than expected (i.e. when there is a *positive* reward prediction error), dopamine neurons increase their firing rates and levels of dopamine release. The dopamine release acts as a learning signal to revise expectations in the positive direction. In contrast, when outcomes are worse than expected (i.e. when there is a *negative* reward prediction error), dopamine neurons are inhibited and release less dopamine. The dip in dopamine acts as a learning signal to revise expectations in the negative direction (Hart et al., 2014; Schultz et al., 1997). Artificially raising dopamine levels buffers the brain from negative prediction errors and interferes with learning from worse-than-expected information, resulting in an exaggerated optimism bias.

Unrealistic optimism is often attributed to cognitive and motivational biases in how people reason about risks and probabilities (Shepperd et al., 2015). For example, when told that one's estimate of developing cancer is lower than the expected risk of someone their age, an individual might compare themselves favorably to the average person their age and conclude that the average estimates do not apply to them. These explanations invoke "high-level" cognitive functions such as reasoning and social comparison. A study, however, suggests that unrealistic optimism might also arise from "low-level" asymmetries inherent to learning mechanisms in the brain (Lefebvre et al., 2017). In this study, participants learned to choose between affectively neutral visual stimuli for monetary reward. Even in this simple scenario where the stimuli had no *a priori* desirability or undesirability, participants updated their beliefs more when outcomes were better than expected, relative to when outcomes were worse than expected. The asymmetric updating was associated with stronger encoding of positive prediction errors in areas of the brain that encoding reward value, including the striatum and the ventromedial prefrontal cortex. Thus, domain-general biases in how the brain processes positive and negative information could contribute to unrealistic optimism. Of course, these results do not preclude a role of high-level motivational and cognitive biases. Examining the extent to which motivated reasoning interacts with biases in reward learning would be a promising direction for future studies of unrealistic optimism.

GROUP DYNAMICS

Our evolutionary need to belong to a group drives us to draw boundaries between in-group and out-group. Across various social contexts, people tend to favor in-group members over out-group members. Intergroup biases have been documented in perception, empathy, and decision-making. Studies using neuroscience tool have begun to identify the brain regions associated with intergroup biases (Cikara & Van Bavel, 2014; Molenberghs & Louis, 2018). In this section, we review early and current work on how the brain sees "us" and "them", as well as the consequences of intergroup biases on empathy and intergroup helping.

Amygdala Responses To In-Group And Out-Group Faces

Early work on the neuroscience of intergroup perception focused primarily on the role of amygdala when viewing in-group and out-group participants, particular in the context of racial groups in the United States. These studies have generally found stronger amygdala activity when viewing faces from a racial outgroup (other-race faces) than when viewing faces from the racial ingroup (same-race faces) (e.g., Hart et al., 2000; Richeson et al., 2008; Ronquillo et al., 2007), an effect that correlates with implicit behavioral measures of racial bias (Cunningham et al., 2004; Phelps et al., 2000). While this effect tends to be reliable in studies with White participants, results from studies that include both White and Black participants have been less consistent – in some cases, amygdala activity in both groups was higher when viewing outgroup faces (e.g., Hart et al., 2000), whereas in other cases, amygdala activity in both groups was higher when viewing African American faces (e.g., Lieberman et al., 2005). One explanation for these inconsistencies is that intergroup perception of faces may depend on cultural experience of an individual (Cloutier et al., 2014), as well as situational goals (Wheeler & Fiske, 2005).

Historically, the amygdala has been associated with the processing of fear and threat-related stimuli (LeDoux, 2000). Thus, the heightened amygdala response to out-group faces was initially interpreted as reflecting negative attitudes towards the out-group. More recent perspectives have emphasized a broader role of the amygdala in processing motivationally salient information (Cunningham & Brosch, 2012; Pessoa & Adolphs, 2010; Pourtois et al., 2013). For example, extremely positive, novel and task-relevant stimuli also have been shown to activate the amygdala. Therefore, stronger amygdala activation to other-race faces may not be due to inherently negative views of the other race, but may instead reflect the salience of other-race faces in contexts where race is salient. In an fMRI study that supports this interpretation, Van Bavel and colleagues (2008) assigned participants to one of two arbitrary groups. As group membership was orthogonal to race, race became a less salient feature for social categorization. Indeed, amygdala responses were not different when participants viewed same-race and other-race faces. Amygdala responses distinguished between in-group faces and out-group faces, even though the groups were newly formed with membership arbitrarily determined. Notably, amygdala activity was *higher* when participants viewed in-group faces than out-group faces. As there was no reason for participants to be prejudiced against in-group members, the authors argued that the heightened amygdala response was unlikely to reflect threat processing, but rather the motivational relevance of in-group faces, which would be the primary driver of amygdala activity in the absence of other salient cues.

Whether the amygdala response to out-group faces reflects prejudice or salience remains a hotly-debated question (Chekroud et al., 2014; Chiao et al., 2008; Santos et al., 2011). What is less uncertain is that the amygdala is sensitive to group membership. In exploring the role of the amygdala in intergroup perception, social neuroscientists broaden the view on amygdala function. As we learn more about the different contextual factors that modulate amygdala responses, we will continue to refine our theories about how the amygdala plays a role in shaping how people see “us” and “them.”

Group Membership Modulates Neural Responses In The Fusiform Face Area

A second line of work has focused on neural responses in the fusiform face area (FFA) to in-group and out-group faces. The FFA is a brain region specialized for the visual processing of faces, and FFA activation in response to viewing faces is one of the most robust findings in cognitive neuroscience (Kanwisher et al., 1997; Kanwisher & Yovel, 2006). In contrast to the amygdala, where activity tends

to be higher when viewing other-race faces than same-race faces, activity in the FFA tends to higher when viewing same-race faces than other-race faces. In an influential study, Golby et al. (2001) used fMRI to scan White and African-American participants as they viewed photographs of White and African-American individuals. Across both groups of participants, FFA activity was higher when viewing same-race faces than other-race faces. Feng et al. (2011) replicated this effect in a sample of Chinese participants viewing White and Chinese faces. Greater FFA activation for same-race faces correlated with better memory of same-race faces, suggesting that the heightened activation to same-race faces is associated with enhanced encoding of same-race individuals relative to those not of the same race.

When participants were arbitrarily assigned to novel groups where membership was orthogonal to race, they exhibited stronger activity when viewing in-group faces than when viewing out-group faces (Van Bavel et al., 2008). FFA activity did not distinguish the race of the viewed face, indicating that FFA was sensitive to the distinction between in-group and out-group faces rather than race per se. In the absence of other group membership cues, people default to categorizing others based on their race. However, if an alternative group identity is salient, participants perceive others as in-group or out-group based on the more salient identity cue. In a follow-up study, the researchers had participants view in-group faces, out-group faces, and a group of faces that did not belong to any particular group. The unaffiliated faces served as a control condition to which the in-group and out-group conditions could be compared. FFA activity was higher when viewing in-group faces than when viewing the unaffiliated faces, but not different when viewing out-group faces than when viewing the unaffiliated faces (Van Bavel & Cunningham, 2011). These results suggest that the heightened FFA activity reflects enhanced processing of in-group faces rather than diminished processing of out-group faces.

Theoretical models of social categorization propose that people tend to think categorically about out-group members. In contrast, when perceiving in-group members, people are more likely to focus on individuating features (Fiske & Neuberg, 1990; Hugenberg et al., 2010). As a result, people tend to think of out-group members as similar to one another, whereas each individual in-group member is perceived as unique. An fMRI study demonstrated that this "out-group homogeneity" effect may have its neural underpinnings in differences in perceptual sensitivity of the FFA to in-group and out-group faces (Hughes et al., 2019). The researchers capitalized on a phenomenon referred to as "neural adaptation" or "repetition suppression," where the repeated presentation of the same or similar stimuli results in attenuated neural responses. That is, when the same or similar stimuli are presented consecutively, the brain habituates such that the neural activation to the subsequent presentations is diminished relative to if the stimuli were presented on its own. The researchers reasoned that if people indeed perceived outgroup members as more similar to one another, there would be stronger habituation to consecutively presented outgroup faces, as would be the case if they were presented with the same face repeatedly.

Using fMRI, Hughes et al. measured the neural response of White participants to blocks of White and Black faces digitally altered to vary systematically on visual similarity. They found that activity in the FFA habituated less to same-race faces than other-race faces, suggesting that the FFA responded to same-race faces as unique, whereas other-race faces were seen as homogenous. Furthermore, the FFA was more sensitive to physical differences in same-race faces than in other-race faces, such that a smaller degree in perceptual difference between consecutively presented faces was sufficient to release the FFA from habituation. These results suggest the FFA is narrowly tuned to distinguish between in-group members, but more broadly tuned to category membership

for out-group members. Future work can use a similar experimental approach to assess if these tuning properties are ingrained by experience, or if they can be modulated by motivational goals.

Intergroup Biases In Empathy

Empathy refers to the ability to share and understand the affective and mental states of others (Decety, 2015). Empathic responses to the suffering of others facilitate prosocial actions to alleviate that suffering. However, these responses are not automatic. Rather, they are modulated by multiple social factors, especially group dynamics (Fourie, Subramoney, & Gobodo-Madikizela, 2017). Reduced empathy for the out-group can thus lead to a decreased propensity to act prosocially towards out-group members. In an early fMRI study of how group membership affects empathic responses, Xu and colleagues (2009) presented Chinese and White participants with video clips of Chinese and White individuals experiencing painful (i.e. pricked by a needle) and non-painful (i.e. touched by a cotton Q-tip) stimulation. Viewing an ingroup member experience painful stimulation was associated with greater activity in the dorsal anterior cingulate cortex (dACC), an effect that was not observed when viewing an outgroup member experience the same painful stimulation.

In a study that utilized a similar experimental paradigm, Contreras-Huerta et al. (2013) assigned White participants to two groups containing a mix of Chinese and White individuals. The participants then viewed video clips of the in-group and out-group members receiving painful or non-painful stimulation. Even though implicit and explicit measures indicated that participants identified with their assigned group, there was no effect of assigned group membership on neural responses to observing painful stimulation. In contrast, activity in the dACC, anterior insula (AI) and somatosensory cortex was higher when observing same-race individuals experience pain than when observing other-race individuals experience pain. Of course, these results do not preclude the possibility that a more involved procedure of inducing group identification beyond a minimal group paradigm would result in intergroup biases in empathic neural responses to others' pain. Nevertheless, they suggest that, unlike the amygdala and FFA response ingroup and outgroup faces, intergroup empathic neural responses to the pain of others may be less malleable to social context.

The dACC, AI, and somatosensory cortex are activated when one experiences pain, as well as when one observes someone else experience pain (Corradi-Dell'Acqua et al., 2011; Jackson et al., 2005). The overlap in brain regions associated with experienced and observed pain has led to the proposal that observing someone else experience pain is aversive, and can motivate behaviors to alleviate the pain of others (Zaki et al., 2016). As a result, reduced empathic responses to the pain of outgroup members can have consequences for intergroup helping. In a test of this idea, Hein and colleagues (2010) had soccer fans watch a fan of their favorite team (i.e. an ingroup member) and a fan of a rival team (i.e. an outgroup member) experience pain. The participants when then given opportunities to reduce the pain of the others by enduring the physical pain themselves. AI activity was higher when participants viewed the ingroup member experience pain than when they viewed the outgroup member experience pain. Participants helped the ingroup member more than the outgroup member, and the degree of helping was associated with stronger activations in the AI. Together, these results suggest that negative affect associated with viewing an ingroup member experience pain can motivate costly helping, and the reduction in this empathic response for outgroup members reduces intergroup helping. In addition, the authors found that the activity in the nucleus accumbens, a region associated with reward and positive affect, was associated with *not* helping the outgroup member, suggesting that a schadenfreude-like signal is also involved in reducing the tendency of intergroup helping.

Altogether, these studies consistently illustrate reduced empathy for outgroup suffering, and present a bleak picture of intergroup relations. Are there interventions that can counteract these automatic tendencies? One study found that neural responses to the suffering expressed by other-race individuals can change over time and is mitigated by familiarity of contact with the outgroup. In this study, the authors recruited Chinese students who had arrived in Australia between six months to five years prior to the experiment and assessed their level of contact with other ethnic groups across various contexts (Cao, Contreras-Huerta, MacFayden, & Cunnington, 2015). During fMRI scanning, participants were shown videos of own-race/other-race individuals, as well as own-group/other-group individuals, expressing pain. The typical in-group bias in neural responses to observed pain was evident, whereby neural activation was greater for pain in own-race compared to other-race people. Critically, the racial bias in neural responses to observed pain decreased significantly with increasing level of self-reported contact with other other-race individuals, suggesting that intergroup contact could help reduce intergroup biases in empathy. In an fMRI study, Hein et al. (2016) set out to test the hypothesis of whether positive experiences with outgroup members can reduce the empathy gap towards the outgroup. Participants encountered either a member of an ethnic out-group (experimental condition) or ethnic in-group (control condition) who gave up resources to save the participant from a painful electric shock. In a later part of the experiment, participants then viewed different members of the ethnic ingroup and outgroup experience pain. Participants in the control condition exhibited the same pattern of results in the earlier studies, where AI response was greater when observing the pain of ingroup members than outgroup members. Participants in the experimental condition exhibited the opposite pattern of results where the AI response was greater when observing the pain of outgroup members than ingroup members. These participants also reported greater empathy for the outgroup after the experiment. While future work is needed to examine if the enhanced AI response to outgroup members would translate to increased intergroup helping, these results nevertheless suggest that positive interactions with the outgroup shape neural responses to outgroup suffering and can reduce the empathy gap between groups.

SOCIAL INFLUENCE

Our decisions are often influenced by the opinions, actions and decisions of others (Turner, 1991). For example, we conform to group norms by changing our perceptions, opinions and behavior to be better aligned with the perceptions, opinions and behavior of our peers; we often seek out and utilize advice when making decisions; and we consider the preferences of others when making decisions in a group. Over the last few decades, there has been a large and growing interest in understanding the neurocognitive correlates of social influence (Izuma, 2013, 2017; Stallen & Sanfey, 2015). Here, we review some of the landmark studies on this topic and examine how they contribute to our understanding of the mechanisms underlying social influence.

Neural Correlates Of Conformity

In an early fMRI study examining the neural correlates of conformity, Klucharev et al. (2009) presented participants with a series of faces and had them rate the attractiveness of each face. After each rating, participants then were shown the purported average ratings of other raters. Unbeknownst to the participants, the average ratings were systematically manipulated to be different from the participant's own ratings on two thirds of the trials. Then, participants were

presented with the faces a second time and re-rated the faces. Consistent with prior behavioral work on conformity, participants' second ratings shifted in the direction of the average rating.

The researchers then examined neural responses when participants first were informed of the average rating to identify brain activity that predicted subsequent conformity. Relative to trials where the average rating was consistent with the participant's rating, trials where the average rating diverged from the participant's rating was associated with an increase in activity in the posterior medial prefrontal cortex (pMPFC), an area of the brain associated with cognitive control. In addition, there was a decrease in activity in the ventral striatum, an area associated with reward processing. Both the activation in the pMPFC and deactivation in the ventral striatum were associated with the extent to which participants later adjusted their ratings to more closely align with the average rating of the group. As discussed in earlier sections of this chapter, the ventral striatum is thought to encode a reward prediction error signal that encodes whether an outcome is better or worse than expected. Thus, the deactivation in the striatum in response to a discrepancy between self and group rating is consistent with the interpretation that deviating from the group norm is aversive (i.e. a negative outcome). This negative reinforcement signal is then conveyed to the pMPFC, which in turn adjusts participants' ratings towards the group norm. In a follow-up study, the researchers used transcranial magnetic stimulation (TMS) pulses to transiently perturb pMPFC function prior to participants performing the task (Klucharev et al., 2011). Relative to participants who did not receive TMS and participants who received TMS to a different region of the brain, participants who received TMS to the pMPFC conformed significantly less to the average judgments. These results complement the earlier fMRI study and provide causal evidence of the involvement of the pMPFC in mediating conformity.

Dissociating Genuine Attitude Change From Public Compliance

It can often be difficult to ascertain whether apparent conformity to group norms reflects public compliance due to social pressures, or a true modification of attitudes arising from internalizing the opinions of the group. In the former case, people may publicly report their attitudes as consistent with group norms, but those reports might not reflect the attitudes they hold. Using a similar paradigm as the earlier studies by Klucharev et al. (2009; 2011), Zaki et al. (2011) demonstrated how fMRI can be used to dissociate true attitude change from public compliance. Participants rated a series of faces based on attractiveness, and then were told the average rating given by their peers. The average ratings of a subset of faces were manipulated to be higher than the rating given by the participant ("peers-higher faces"), whereas for a different subset of faces the average ratings were manipulated to be lower than the rating given by the participant ("peers-lower faces"). Following the manipulations, participants rated the faces a second time. Replicating the earlier results of Klucharev et al. (2009), participants changed their ratings to conform to those of their peers – giving the peers-higher faces higher ratings and peers-lower faces lower ratings than they did initially.

While Klucharev et al. scanned participants when they were told the average rating of their peers, Zaki et al. scanned participants during the second ratings. The authors reasoned if conformity reflects true attitude change, the effect should be reflected in the neural response to these images. Building on prior work indicating that activity in the ventral striatum and orbitofrontal cortex- areas of the brain that are associated with encoding reward value- correlate with facial attractiveness, the authors examined how peer ratings of the faces affected the neural activity in these regions at the time of the second rating.

Activity in both the ventral striatum and orbitofrontal cortex was higher when viewing faces that peers rated as more attractive ("peers-higher faces") than when viewing faces that peers rated as less attractive. As the faces were matched on initial attractiveness, these differences in neural activity were the result of having been exposed to the peer ratings. Furthermore, the directionality of the effects is consistent with the interpretation that participants perceived the peers-higher faces as more attractive and peers-lower faces as less attractive than they did originally. Thus, these results suggest that social influence changed participants' attitudes, and that their conformity was not merely public compliance to group norms. Certainly, the results do not imply that conformity *always* reflects true attitude change, but they highlight how fMRI can be used as a tool to detect it when it does.

Social Preference Moderates Conformity Effects

According to balance theory, people strive towards maintaining "cognitive balance" in their attitudes and beliefs (Heider, 1946). Having a different preference from a liked other or having a similar preference to a disliked other creates cognitive imbalance, which motivates an individual to change their preference to restore cognitive balance. As such, people tend to shift their preferences closer to the preferences of liked others and away from the preferences of disliked others. Izuma and Adolphs (2013) used fMRI to examine the neural correlates of cognitive imbalance and how cognitive imbalance drives attitude change. In their study, participated rated their preferences of T-shirt designs and were then told the average preference of a liked group (fellow students of their university) and a disliked group (sex offenders). In line with the predictions of balance theory, participants shifted their preferences to align with the preferences of the liked group and away from the preferences of the disliked group.

Activity in the pMPFC tracked the degree of cognitive imbalance on each trial. Specifically, the pMPFC was more active when the liked group also disliked the T-shirt the participant liked, as well as when the disliked group liked the T-shirt the participant liked. The magnitude of pMPFC activation was associated with the degree of subsequent attitude change. The involvement of the pMPFC in both conforming to liked others and distancing from disliked others suggest that the two processes draw on similar neural mechanisms. Notably, the pMPFC has also been shown to correlate with cognitive dissonance when making difficult decisions for oneself (Izuma et al., 2010). Thus, the pMPFC may encode a general cognitive inconsistency signal and various forms of cognitive inconsistencies posited by different social psychological theories might share common neurobiological roots.

Learning From Advice

When making decisions, people often seek out and utilize advice from others. Seeking out expert advice allows people to capitalize on the expertise and knowledge of others who know more than they do. In doing so, people can avoid costly mistakes and make better decisions. The propensity to seek out, take, and give advice is thought to be crucial to the growth and evolution of human culture (Boyd & Richerson, 1988). To study the neural mechanisms underlying advice-taking, Biele et al. (2011) had a group of participants perform a decision-making task where they had to learn from trial-and-error to choose from one of four card decks associated with different gains and losses. These participants then wrote down their recommendation of which card deck to choose,

and their recommendations were communicated to a second group of participants who performed the same task in the scanner.

The advice led the second group of participants to select the most rewarding deck in a fewer number of trials. As the reward associated with the decks was probabilistic, drawing from the most rewarding deck can at times yield losses. As such, participants sometimes switched away from the recommended deck to explore the other decks. Nevertheless, participants were, in general, more likely to draw from the recommended decks than other decks. Participants' choice behavior was best explained by a computational model that implements an "outcome-bonus" for recommended decks. In this model, receiving rewards from the recommended decks reinforced choice *more* than receiving the same amount of reward from non-recommended decks, while losses from the recommended decks inhibited choice *less* than losses of the same amount from the non-recommended decks. These modeling results were mirrored in the brain, where choosing the recommended deck was associated with higher activity in the ventral striatum, irrespective of the outcome. Together, the modeling and neural results provide converging evidence that following advice might be inherently rewarding.

When advice is sound and well-intentioned, following advice can be strategic and optimal. However, in cases when advice is itself misinformed or malicious, following advice can instead lead to bad decisions and costly mistakes. In these cases, the "outcome-bonus" associated with following advice can lead to people persevering in following bad advice (Doll et al., 2009; Leong & Zaki, 2018). In a decision-making task where participants were given advice that is on average inaccurate, ventral striatum response to positive outcomes when following the advice was associated with the tendency to perseverate when following the bad advice, even though participants would have performed better if they had ignored the advice (Staudinger & Büchel, 2013). This effect was particularly strong if the positive outcomes occurred early on in the task (i.e. a primacy effect), suggesting that the initial positive feedback solidifies the perceived efficacy of the advice, and once solidified, the perceived efficacy of the advice becomes resistant to revision even in the face of disconfirming evidence. These results provide a possible explanation for why investors often follow the investment advice of stock pundits who are no better than chance at predicting the market, and consumers adopt questionable health remedies recommended by medical talk shows.

Several studies have examined the neural mechanisms underlying how people learn whose advice to follow and whose advice to ignore (Behrens et al., 2008; Boorman et al., 2013). For example, Boorman et al. (2013) had participants make multiple decisions on whether or not to invest in a stock over time. The performance of the stock followed predictable patterns, and participants could base their decision on the past performance of the stock. On each trial, participants were provided with the advice of a financial advisor. Some of the financial advisors were accurate, predicting the performance of the stock with above chance accuracy, whereas other advisors were inaccurate, predicting the performance of the stock with below chance accuracy. Participants had to learn which advisors were accurate and which advisors were inaccurate through trial-and-error.

Participants combined information about the past performance of the stock with advice they received from the advisors to make their decisions. Importantly, they learned to follow the advice of accurate advisors and ignore or go against the advice of inaccurate advisors. Analyses using a computational model suggest that when estimating the expertise of an advisor, participants not only considered the outcomes of the advisor's prediction (i.e. whether the advisor was accurate or inaccurate), but also the extent to which the advisor's advice was consistent with the participant's beliefs about the stock. When an advisor was inaccurate, participants penalized them less when

their advice was consistent with the participant's own beliefs about the stock, and more when their advice was inconsistent with the participant's beliefs, suggesting that participants considered *both* the "quality" of the advice (i.e. whether the advice was sensible) and the actual outcomes.

Learning about the advisors' expertise was associated with activity in the mentalizing network, including the MPFC, TPJ and precuneus. When the advisors were replaced with computer algorithms, these same regions were associated with learning about the algorithm's ability to predict the stock. Whether the involvement of these regions in learning about the expertise of both human advisors and computer algorithms reflect a domain-general system of outcome-contingency learning or the anthropomorphizing of computer algorithms as social agents is unclear. In this study, the researchers did not examine if there was a systematic bias towards following advice. A behavioral study using a similar paradigm showed that while participants were able to discern the relative accuracy of accurate and inaccurate advisors, participants were nevertheless biased towards following advice more than they should (Leong & Zaki, 2018). Thus, while people are capable of discerning good from bad advice, they require a disproportionate amount of evidence to overcome their natural tendency to discount bad advice.

COOPERATION AND SOCIAL PREFERENCES

While cooperation is common across many species and levels of biological organizations, humans are exceptional at cooperation because it extends beyond close kin to include total strangers and therefore occurs on a much larger scale than other species. Gene–culture coevolution is responsible for human social preferences and accounts for our remarkable capacity to cooperate in large groups (Gintis, 2011). Self-interest, however, is still an important motivator. Both prosocial and egoistic motivations shape decision-making.

Evolutionary theory provides a theoretical framework to explain why cooperative behaviors emerge in nature and assumes that these abilities contributed to the reproductive success of our ancestors. The general idea is that humans are not only selfish but also many are intrinsically motivated by principles of fairness and reciprocity. This is entirely consistent with the fact that prosocial behavior is a product of natural selection and contributes to the reproductive success of our species. There are two explanations for the function of sociability: the ultimate explanation and the proximal explanation. The ultimate explanation suggests that sociability improves the chances of survival and reproduction through the strengthening of social cohesion and cooperation. The proximal explanation maintains that sociability is partly underpinned by intuitive mechanisms, or non-conscious motivations, which are relayed and reinforced by complex cognitive mechanisms and require an efficient brain. It is important to note that we are not aware of the ultimate causes of our adaptations. The distinction between ultimate (evolutionary) and proximal (psychological) levels is important because it allows us to overcome the problems of rational choice theories in the context of prosociality and cooperation.

A behavior is cooperative if it provides a benefit to both the actor and the recipient and if it has evolved at least partially because of this benefit. Cooperation occurs across all scales and levels of biological organization throughout the natural world. For instance, genes cooperate within the genome, organelles cooperate to form eukaryotic cells, bacteria cooperate to overcome host defenses, and humans and social insects cooperate to build societies (Ågren, Davies, & Foster, 2019). Thus, cooperation emerges as a solution (i.e., its adaptive value) during evolution in multiple species and levels, from cells to societies (West, Griffin, & Gardner, 2007).

Cooperation among humans has resulted in the success of this recent species in colonizing the entire planet in a few tens of thousands of years. Humans built pyramids, nuclear power plants, waged war, created social institutions, and raised taxes for the common good to note a few examples. Humans are unique in that they extend cooperation to a wide range of other conspecifics beyond kin, such as to reciprocating partners and even strangers.

Across species, the evolution of cooperation is based on the assortment principle, which states that cooperation can evolve if cooperators selectively direct benefits to other cooperators (Apicella & Silk, 2019). Several mechanisms have evolved to suppress selfish behavior and promote cooperation in humans, including the reinforcement of social norms. Note that enforcement, which is an action that evolves at least in part to reduce selfish behavior within a cooperative system or alliance, is not limited to homo sapiens. In reality, such mechanisms exist at all biological scales, from conflict mediation in mitochondrial–nuclear interactions for eukaryotic cells, to enforcement of cooperation in microbes, to policing in animal societies (Ågren, Davies, & Foster, 2019).

Sapiens As An Ultrasocial Species

Cooperation among unrelated individuals, who do not share direct genetic interest in offspring, is ubiquitous in human societies, and the propensity to cooperate can be observed early in ontogeny (Decety, Steinbeis, & Cowell (2021). Children as young as 18 months of age seem intrinsically motivated to help others to achieve their goals in a variety of different situations without external reinforcement (Tomasello, 2020). Moreover, before their second birthday, toddlers develop equity-based expectations of fairness in which individuals are expected to receive the treatment they deserve in specific situations, like an equitable share of windfall resources. More specifically, toddlers expect third-party to receive a reward proportionate with their efforts or a punishment befitting their transgressions (Tin et al., 2020). Later in development, children adjust their prosociality with group preferences and social norms (Huppert et al., 2019).

Other-Oriented Concerns

Most economic models are based on the self-interest hypothesis, which assumes that individuals are exclusively motivated by their material self-interest. Evidence from behavioral studies suggests otherwise, showing that people are strongly motivated by other-regarding preferences, care about social norms, and punish those who exploit the cooperative behavior of others. Therefore, overall concerns for fairness and reciprocity cannot be ignored in economic decisions (Fehr & Schmidt, 2006). The cognitive processes and underlying neural circuits involved in cooperation involve the interplay of multiple systems including the ability to think about mental states (one's own and others'), emotional awareness, reward circuits, valuation, and integration of social norms. It has been proposed that humans are unique in their abilities to represent shared intentions and goals and to self-regulate in order to comply with and enforce cooperative norms (Zoh, Chang, & Crockett, 2022). Such abilities, according to the authors of that review, rely on three intertwined neural networks: 1) a medial one that values outcomes for self and others; 2) a lateral prefrontal network that guides cooperative goal pursuit by modulating value representations in the context of local norms; and 3) an anterior prefrontal network that serves the uniquely human abilities to reflect on one's behavior, commit to shared social contracts, and arbitrate between cooperative norms across diverse social contexts (Zoh et al., 2022).

Humans not only show a strong aversion to inequity in behavioral economics games, as reflected by responses in the anterior insula, anterior cingulate cortex, and dorsolateral prefrontal cortex, but also exhibit similar emotions when others are treated unfairly (Tricomi, & Sullivan-Toole, 2015). The advantageous inequality aversion, or paying at a personal cost to achieve equal reward distribution, represents a unique feature of human cognition. Both types of unfairness similarly activate the striatum as well as the orbitofrontal cortex and anterior insula, regions associated with representing subjective value and aversion, respectively (Yu, Calder, & Mobbs, 2014). In 4-year-old children, electrophysiological responses distinguish very rapidly in time (around 275 ms) as a result of witnessing equal from unequal third-party resource distributions (Cowell, Sommerville, & Decety, 2019). Around 500 ms, the electrophysiological response distinction predicts children's decisions in a distributive justice game where the recipient varies according to wealth, merit, or need. This late electrophysiological component reflects the engagement of attentional resources and the processing in short-term memory, allowing a conscious representation of the choices they will make.

At the same time, neurohormones also play a role in modulating preferences for cooperation (Crockett & Rini, 2015). For instance, enhancing serotonin function increases cooperation in the prisoner's dilemma whereas its depletion decreases it. Neuropeptide oxytocin motivates trust and cooperation with reciprocating others (Rilling et al., 2012) and in-group members (Israel, Weisel, Ebstein, & Bornstein, 2012). It is a nine-amino acid produced primarily in the hypothalamus, and then shuttled to the pituitary for peripheral release or projected to various brain regions, including the amygdala and reward- and reinforcement pathways. Homologs of oxytocin have existed for at least 700 million years, and have been identified in such diverse organisms as hydra, worms, insects, and all vertebrates (Donaldson & Young, 2008). Oxytocin's role in facilitating species-typical social and reproductive behaviors is evolutionarily conserved, although the specific behaviors that it regulates are quite diverse. In humans, oxytocin influences social cognition in many ways:

- It facilitates parental attachment, social synchrony, bonding, and caring (Michalska et al., 2014; Szymanska et al., 2017).
- It promotes trust and the willingness to accept social risks arising through interpersonal interactions (Kosfeld et al., 2005; Mikolajczak et al., 2010).
- It increases both parochial and universal altruism (Israel et al., 2012).
- It motivates in-group favoritism and out-group derogation (De Dreu et al., 2011).
- It improves empathy (Shamay-Tsoory et al., 2013) and theory of mind (Domes et al., 2007).
- It is involved in defense-motivated non-cooperation towards rivaling out-groups (DeDreu et al., 2012).
- It increases attention and arousal to both social and non-social stimuli that are motivationally relevant (Eckstein et al., 2019).

Overall, oxytocin seems to facilitate the processing of social and non-social sensory cues that are crucial for survival, depending on various contextual factors (Quintana & Guastella, 2020).

Lastly, reputation plays a key role in promoting cooperation. Individuals who have prosocial reputations are trusted more, have more influence, and are more likely to be selected as exchange partners (Simpson & Willer, 2015). Within a social network, reputation is both a signal to attract valued partners and a reward at the individual level. In one study, participants were asked to give any amount of money to charities or to keep the money, when they were either alone or believed to be observed by others (Izuma, Saito, & Sadato, 2010). Being observed by others enhanced the donations, and increased activity in the ventral striatum before decision-making, showing that

reward processing motivates decision-making. Relatedly, darkness induces a sense of anonymity, and as a result, can reduce concerns about one's reputation. For instance, when participants were given an opportunity to cheat on a test and thus gain higher monetary rewards, less than 25% did so when the room was brightly lit but more than 50% cheated when the room was less well lit (Zhong et al., 2010).

Cooperation Is Inherently Rewarding

Endogenous rewards significantly increase cooperation in public goods games, and the trend of cooperation increases with rewards. This is what the results of an experiment conducted by Yang, Zhang, Charness, Li, & Bien (2018) demonstrate. The authors found that players who interacted anonymously with the same three other players throughout are much more likely to reward higher contributors than lower ones, regardless of their own contributions, and that most of the participants in the experiments are conditional cooperators, who use the average contribution level in the group as a reference. Furthermore, study participants reward high contributors and receive rewards, in turn, themselves when they make large contributions, thus cultivating a virtuous cycle. Several studies have documented increased neurodynamic signal in the striatum and ventromedial prefrontal cortex when participants engaged in mutually cooperative activities and interpreted this pattern as reflecting activation of the reward circuit. For instance, one study scanned participants while playing an iterated Prisoner's Dilemma Game with another person (Rilling et al., 2002). Mutual cooperation is associated with increased signal in areas involved in reward processing, including the nucleus accumbens, caudate nucleus, and ventromedial prefrontal cortex. That led the authors to propose that activation of this neural network reinforces reciprocal altruism by motivating participants to resist the temptation to selfishly accept but not reciprocate favors. Similar results were found in another early fMRI study when participants played a computer game in cooperation with another person in comparison to playing in competition against someone else (Decety et al., 2004).

Even during a single interaction with an on-line stranger where people could lose money, participants exhibit a preference for cooperation (Sun et al., 2016). This decision is associated with a rewarding experience from a mutually cooperative social interaction. Such cooperation was associated with activation of the ventral striatum and the orbitofrontal cortex. In the study, participants played a series of one-shot prisoner's dilemma games in the loss and gain contexts. They reported a greater level of positive emotions for mutual cooperation in both contexts, but they cooperated more in the gain compared to in the loss context, which was paralleled with a significant increase in the orbitofrontal cortex and ventral striatum.

Altruistic Punishment

Studies in experimental economics and game theory have consistently documented that people are rational in their decisions to collaborate or not (Fehr & Gachter, 2000). People distinguish short-term from long-term interactions and cooperate much more in the latter case. They also adjust their level of altruistic contribution rationally according to costs and benefits. When costly punishment is permitted, cooperation doesn't deteriorate in public goods games.

Altruistic punishment, when individuals punish norm violators at a cost to themselves, is a powerful mechanism to stabilize steady levels of cooperation among genetically unrelated

individuals and to solve the "free rider" problem in public goods game. Importantly, individuals are willing to enforce norms not just when they are the victim of a transgression, but also when they are unaffected by the transgression, which has group-level benefits in maintaining social norms.

People feel bad when they witness norm-breakers go unpunished and derive satisfaction from their punishment. This is examined using trust games, in which one player's profits are increased if they invest money and a second player cooperates to return half of the proceeds, but that same player's profits are decreased if the second player defects and keeps the investment. In these games, punishing non-cooperators at a cost activates the caudate nucleus, which is part of the reward circuit (Quervain et al., 2004). Furthermore, a positive correlation was found between the signal increase in that region and the investments in punishment. Another reward-processing region, the nucleus accumbens, similarly showed punishment-related activation during a third-party punishment condition in a one-shot economic exchange game, which warrants strong reciprocity (Strobel et al., 2011). A different study examined the neural circuitry underlying social norm compliance by comparing a condition during which violations of a fairness norm is possible with a condition during which punishment is impossible (Spitzer et al., 2007). Behavioral results demonstrate that the social punishment threat was very effective in inducing subjects to conform to the fairness norm. Activation in the dorsolateral cortex, orbitofrontal cortex (implicated with the evaluation of punishing cues and inhibition of prepotent action tendencies), and caudate nucleus were predictive of individual participants' compliance (Du & Chang, 2015 for a review on the neural components of altruistic punishment).

Individual variation in costly punishment behavior is modulated by serotonin, a neurotransmitter that influences social decision-making within the flight-or-flight response (Lee & Goto, 2018). One study that combined manipulations of serotonin during an Ultimatum Game found that participants were more likely to punish those who treated them unfairly, and were slower to accept fair exchanges (Crockett et al., 2013). Depletion of serotonin simultaneously reduced ventral striatal responses to fairness and increased dorsal striatal responses during punishment decisions. The same group of social neuroscientists had also demonstrated that enhancing serotonin function reduces costly punishment by increasing aversion to harming others (Crockett et al., 2010). Some studies showed that testosterone administration causes an increase in punishment in behavioral economics games. For instance, participants who received testosterone punished their opponents more frequently than those administered with a placebo, and higher peripheral testosterone levels were specifically associated with increased punishment of opponents who made unfair offers (Dreher et al., 2016).

Self-Interested Motivation Is Still Alive

People are strongly motivated by fairness and reciprocity and are willing to reward or punish other people at a considerable cost to themselves, but many individuals pursue self-interest if they can maintain the appearance of being fair (Caviola & Faulmüller, 2014). Prosocial behaviors thus decline when people are not visible to other affected parties (Overgaauw et al., 2012). One study manipulated fairness for the self and other in a three-party economic game, and employed both functional MRI and high-density electroencephalography to characterize the spatiotemporal neural dynamics sensitive to the payoffs to oneself or to a powerless anonymous observer (Yoder & Decety, 2020). A greater hemodynamic response was detected in the anterior insula and a greater mid-frontal amplitude were detected for unfairness towards the self when compared with a neutral observer. Furthermore, across the fMRI and the EEG measures, participants were more sensitive to

the plight of the neutral observer when they themselves received at least one-third of the initial endowment. Finally, a spatial multivariate pattern analysis identified the right anterior insula and dorsolateral prefrontal cortex as regions capable of decoding fairness, but only for the self. This result provides additional support for a "me first" mode of processing.

Large-scale cooperation is a distinctive feature of our species. Sociological factors such as rules, descriptive and injunctive norms, and coercive and reinforcement mechanisms play synergistic roles with our innate propensity for sociability in maintaining the exceptional levels of cooperation that characterize human societies.

NEW FRONTIERS

Social neuroscience is a deeply interdisciplinary enterprise that emphasizes the functions that emerge through the interaction of conspecifics, the neurobiological mechanisms underlying these functions, and the commonality and differences across social species. In doing so, the ultimate goal is to understand the influences and associations between the social and biological levels of organization. As is true with any academic discipline, progress will be spurred by advances in both theory and methodology.

At the theoretical level, an important new frontier is to develop theoretical frameworks that integrate theories from social psychology and cognitive neuroscience. An early example of this is the dynamic interactive theory of person construal proposed by Freeman and Ambady (2011), which combines classic social psychological theories of person perception with neuroscience models of visual perception. The theory incorporates key insights from social psychology, while also capitalizing on what we know about the visual system. Freeman and colleagues (2018) have since built on this model to make specific neural predictions about where and how social information is represented and used by the brain, and found support for these predictions in neural data. This line of work is a great illustration of how theories in social neuroscience are moving beyond finding neural support for existing social psychological theories, and towards building theories that seek to explain both neural and social psychological phenomena concurrently.

Can Cognitive Processes Be Predicted From Functional Imaging Data?

A notable challenge in theory-building is that most neuroscience data (though not all) are inherently correlational, and as such, it is difficult to avoid reverse inference, that is, the reliance on reasoning backward from patterns of activation to infer the engagement of specific mental processes. Few people would argue with Poldrack (2006) when he pointed out that a specific brain region can be activated by a wide range of tasks and cognitive processes. For example, the amygdala not only responds to fear but more generally to emotional saliency (positive and negative events) relevant for the goals and motivations of the perceiver (Cunningham, & Brosch, 2012). Similarly, the anterior insula is involved in the first-hand experience of nociception and vicarious pain, but also in sexual pleasure. More generally, this region serves a critical role in emotional awareness and interoceptive attention by integrating bottom-up interoceptive signals with top-down predictions (Wang et al., 2019). Likewise, a many-to-one relationship exists between network structure and

network behavior (Marder, Goeritz, & Ottopalik (2015). Thus, one has to be careful when drawing inferences about mental processes from neural activations.

However, the reverse inference cannot be wholly disregarded as a fallacy per se. Rather, the predictive power of reverse inference can even be "decisive"—dependent on the cognitive process of interest, the specific brain region activated, and the task-setting used (Hutzler, 2014). Progress in this domain relies largely on the ability to bridge the gap from MRI to brain structure and to examine the multiple realizability of behavior in brain structure. Recent developments along these lines, such as advanced MRI techniques and advanced statistical methods, provide opportunities for a better conceptual understanding of how multifaceted human behavior emerges from human neural structure (Song et al., 2022).

The development of large-scale automated meta-analytic tools such as neurosynth (https://neurosynth.org/) also allows social neuroscientists to draw stronger reverse inferences (Yarkoni et al., 2011). Specifically, this platform for large automated synthesis of fMRI data applies text-mining methods to a large number of cognitive neuroscience papers to quantify the relationship between a "term" (e.g., "pain") with a brain area. Thus, researchers can obtain probabilistic measures of the likelihood activation in a brain area is associated with a term, the likelihood activation in a brain area is *not* associated with a term, and the likelihood that activation in a brain area is associated with other terms, which then allows for more rigorous reverse inference. For example, if a brain area is strongly associated with "pain" but not with other terms, one can be more confident that activation in that area is associated with pain.

Related, changing norms of data sharing allows researchers to assess the specificity and generalizability of hypothesized brain-behavior relationships across a variety of tasks and datasets. In a new study, Chen and colleagues (2021) identified a multivariate pattern that successfully predicts whether a participant will trust a partner in an interpersonal investment game with 90% accuracy, which also generalizes to a version of task collected in a different country (i.e. a "neural signature" of trust). Importantly, the authors then tested the pattern on eleven additional publicly available datasets and close to 500 participants performing a range of different tasks. The authors found that the pattern could be used to predict beliefs of safety, and was inversely related to negative affect, but was unrelated to reward anticipation, cognitive control, social perception, and self-perception. Thus, their new approach allowed researchers to characterize the specificity and discriminant validity of an identified neural signature, as well as provide theoretical insight in the psychological nature of trust and its relationship with other constructs. The inclusion of close to 500 participants also helps address the common criticism that neuroimaging studies have small sample sizes, and increases the generalizability and reliability of the findings. This analysis was only possible because researchers were willing to share their data on open repositories such as OpenNeuro (https://openneuro.org/) and NeuroVault (https://neurovault.org/).

Opportunities Afforded By New Methods And New Partners

Causal Methods

Another strategy to move beyond reverse inference is to use causal methods that allow drawing stronger inferences about the relationship between psychological processes and brain function. These methods include transcranial magnetic stimulation (TMS), which allows for the transient

perturbation or inhibition of brain function. Another exciting method made possible by new technological advances is the use of real-time neurofeedback to bias participants to modulate their patterns of neural activity, and then assessing the effects of the modulated neural patterns on behavior (Watanabe et al., 2017).

Computational Modeling

The increasing use of computational models in social neuroscience represents another promising approach in uncovering the neural mechanisms underlying social psychological phenomenon (Charpentier & O'Doherty, 2018; Cheong et al., 2017; Hackel & Amodio, 2018). Using computational models, researchers can decompose ongoing task behavior into latent psychological components (e.g., the subjective value of a current choice, the rate of evidence accumulation in an experimental condition), which can then be correlated with neural data to identify the computational role performed by different brain regions (e.g., Hackel et al., 2015; Hutcherson et al., 2015). The models can also be used to estimate individual difference variables that are not directly observable from task behavior. For example, Lockwood and colleagues (2016) scanned participants using fMRI while they learned to choose between different stimuli that were probabilistically associated with rewards for themselves (self-condition) or another person (prosocial-condition). In both conditions, participants were able to learn through trial-and-error to choose the high-reward option. Using a computational model, the researchers identified trial-by-trial learning signals and showed that the subgenual anterior cingulate cortex (sgACC) tracked learning signals only when learning for others. The researchers also used a model to estimate the learning rate for each participant in each condition, and found that participants on average learned faster when earning reward for themselves than when earning reward for others, but this difference was moderated by trait empathy – highly emphatic individuals learned faster for others than for themselves. Furthermore, the sgACC of highly empathic individuals was more responsive to learning signals in the prosocial condition. These results highlight how computational models can be used to better characterize individual differences in neural responses and how they contribute to divergent behavioral outcomes.

In addition to serving as analytical tools, computational models can also reflect formal hypotheses of the mechanism or "algorithm" underlying a social psychological phenomenon (Lockwood et al., 2020). By assessing which model provides the best description of participants' behavior and neural data, researchers can test which of these hypotheses is best supported by empirical evidence. Computational modeling thus allows social neuroscientist to characterize *how* a social psychological process is implemented in the brain, as opposed to merely identifying *where* the process occurs (O'Doherty et al., 2007). There are, however, potential pitfalls associated with the use of computational models in social neuroscience. First, the increasing use of computational models might bias social neuroscientists towards tasks that are more amenable to modeling. This could result in a narrow focus on specific paradigms or experimental designs that are easily quantifiable and can be modeled, while neglecting other complex and dynamic aspects of social cognition that may be difficult to capture using current computational approaches. To avoid falling into this pitfall, social neuroscientists ought to continue using a diverse range of methods and paradigms, while simultaneously developing new computational approaches that capture varied aspects of human cognition. Second, there are an infinite number of possible models, many of which can make similar predictions despite having different underlying architecture and algorithms. How can social neuroscientist determine that they have found the "true" model that describes the human mind? On this note, it is useful to remember the adage "all models are wrong, but some are useful" (Box,

1976). Social neuroscientists ought to always keep in mind that all models are simplified representations of reality. They are useful in uncovering patterns in the data that provide new insights into the underlying mechanisms. However, it is important to not confuse an abstraction of reality with reality itself. Social neuroscientists should be mindful of the limitations of their models and strive to use them as tools for generating insights and hypotheses, rather than as definitive representations of the human mind.

Social Neuroscience In The Wild

New methods for measuring behavioral, physiological and neural activity (e.g., mobile eye-trackers, functional near-infrared spectroscopy) allow for the study of social psychological phenomena "in the wild," with multiple interacting minds. For instance, Báez-Mendoza and colleagues (2021) used wireless intracranial recordings in rhesus macaques during a three-agent group interaction task that taps into the foundation of economic trade (giving, receiving, and reciprocating) in humans. Neurons in dorsomedial prefrontal cortex selectively represent the interaction between specific group members and their reciprocation, retaliation, and past behaviors. These populations of neurons encoded information not only about the decisions of monkeys but also about the directionality of the interactions between them, even when the animals themselves were not involved. This opens a new frontier in social neuroscience, where researchers develop the conceptual and analytical tools to take advantage of these methods to examine social behavior in naturalistic real-world settings.

As social neuroscience moves further "in the wild," it will create new opportunities to study the neural basis of social cognition in applied settings, including in health and education. For example, social neuroscientists are working with teachers to use portable EEG sets to study learning as it takes place in real-world classrooms (Dikker et al., 2017). These applied settings are often an important testbed of social neuroscience theories, where researchers can be more confident of the external validity their work. Furthermore, interfacing directly with society will facilitate the cross-fertilization of ideas and insights. Teachers, doctors, and parents often have domain-expertise that may not be directly accessible to the scientist in the laboratory. We anticipate the line between the academy and society to continue to blur in the field of social neuroscience, which we believe will bring new insight and energy as we continue to explore the neural basis of social phenomena.

Fostering Interdisciplinary Research

Major advances in social neuroscience are increasingly contingent on interdisciplinary research that requires close collaboration between social scientists and neuroscientists. Social neuroscience is inherently interdisciplinary. Methods and approaches from other disciplines can be incorporated to study multiple aspects of our sociality across levels of organization. Often, the most interesting questions lie at the intersection of disciplines such when integrating cognitive neuroscience, mathematical modeling, social psychology, and sociology to examine the reciprocal interactions and emerging functional properties of social networks. Although research in single disciplines will continue to make important contributions to understanding neural and computational mechanisms, collaborating across disciplines, with the goal of integrating data and explanations, should be encouraged. Studying complementary aspects of the same question is not only possible but necessary for a comprehensive understanding of human sociality. It adds value to each of individual discipline and is also highly intellectually rewarding. One challenge to overcome is providing mathematical modeling that is compatible across levels and disciplines. The value of

models stems from the way that they force us to define terms and use consistent categories. This is useful to identify ill-defined concepts or relationships and to find patterns of relationships that are not clearly visible.

REFERENCES

Ågren, J. A., Davies, N. G., & Foster, K. R. (2019). Enforcement is central to the evolution of cooperation. *Nature Ecology & Evolution*, *3*(7), 1018–1029. https://doi.org/10.1038/s41559-019-0907-1

Alexander, R. D. (1974). The evolution of social behavior. *Annual Review of Ecology and Systematics*, *5*(1), 325–383. https://doi.org/10.1146/annurev.es.05.110174.001545

Algoe, S. B. (2019). Positive interpersonal processes. *Current Directions in Psychological Science*, *28*(2), 183–188. https://doi.org/10.1177/0963721419827272

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*(401), 464–477. https://doi.org/10.2307/2234133

Apicella, C. L., & Silk, J. B. (2019). The evolution of human cooperation. *Current Biology*, *29*(11), 447–450. https://doi.org/10.1016/j.cub.2019.03.036

Ashar, Y. K., Andrews-Hanna, J. R., Dimidjian, S., & Wager, T. D. (2017). Empathic care and distress: predictive brain markers and dissociable brain systems. *Neuron*, *94*(6), 1263–1273. https://doi.org/10.1016/j.neuron.2017.05.014

Baek, E. C., Porter, M. A., & Parkinson, C. (2021). Social network analysis for social neuroscientists. *Social Cognitive and Affective Neuroscience*, *16*(8), 883–901. https://doi.org/10.1093/scan/nsaa069

Báez-Mendoza, R., Mastrobattista, E. P., Wang, A. J., & Williams, Z. M. (2021). Social agent identity cells in the prefrontal cortex of interacting groups of primates. *Science*, *374*(6566), 4149. https://doi.org/10.1126/science.abb4149

Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology*, *91*(4), 612–625. https://doi.org/10.1037/0022-3514.91.4.612

Balcetis, E., & Dunning, D. (2010). Wishful seeing: More desired objects are seen as closer. *Psychological Science*, *21*(1), 147–152. https://doi.org/10.1177/0956797609356283

Bangee, M., Harris, R. A., Bridges, N., Rotenberg, K. J., & Qualter, P. (2014). Loneliness and attention to social threat in young adults: Findings from an eye tracker study. *Personality and Individual Differences*, *63*, 16–23. https://doi.org/10.1016/j.paid.2014.01.039

Bartal, I. B.-A., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science*, *334*(6061), 1427–1430. https://doi.org/10.1126/science.1210789

Beer, J. S., & Hughes, B. L. (2010). Neural systems of social comparison and the "above-average" effect. *NeuroImage*, *49*(3), 2671–2679. https://doi.org/10.1016/j.neuroimage.2009.10.075

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249. https://doi.org/10.1038/nature07538

Bellmund, J. L. S., Gärdenfors, P., Moser, E. I., & Doeller, C. F. (2018). Navigating cognition: Spatial codes for human thinking. *Science*, *362*(6415), 6766. https://doi.org/10.1126/science.aat6766

Ben-Ami Bartal, I., Breton, J. M., Sheng, H., Long, K. L., Chen, S., Halliday, A., ... & Kaufer, D. (2021). Neural correlates of ingroup bias for prosociality in rats. *Elife*, *10*, e65582. https://doi.org/10.7554/eLife.65582

Ben-Ami Bartal, I., Shan, H., Molasky, N. M. R., Murray, T. M., Williams, J. Z., Decety, J., & Mason, P. (2016). Anxiolytic treatment impairs helping behavior in rats. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00850. https://doi.org/10.3389/fpsyg.2016.00850

Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science, 22*(7), 891-893. https://doi.org/10.1177/0956797611413

Bernhardt, P. C., Dabbs Jr, J. M., Fielden, J. A., & Lutter, C. D. (1998). Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. *Physiology & Behavior*, *65*(1), 59–62. https://doi.org/10.1016/S0031-9384(98)00147-4

Berntson, G. G., & Norman, G. J. (2021). Multilevel analysis: Integrating multiple levels of neurobehavioral systems. *Social Neuroscience*, *16*(1), 18–25. https://doi.org/10.1080/17470919.2021.1874513

Biele, G., Rieskamp, J., Krugel, L. K., & Heekeren, H. R. (2011). The neural basis of following advice. *PLOS Biology*, *9*(6), 1001089. https://doi.org/10.1371/journal.pbio.1001089

Bilek, E., Ruf, M., Schäfer, A., Akdeniz, C., Calhoun, V. D., Schmahl, C., Demanuele, C., Tost, H., Kirsch, P., & Meyer-Lindenberg, A. (2015). Information flow between interacting human brains: Identification, validation, and relationship to social expertise. *Proceedings of the National Academy of Sciences*, *112*(16), 5207–5212. https://doi.org/10.1073/pnas.1421831112

Boorman, E. D., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron*, *80*(6), 1558–1571. https://doi.org/10.1016/j.neuron.2013.10.024

Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone, and winning and losing in human competition. *Hormones and Behavior*, *23*(4), 556–571. https://doi.org/10.1016/0018-506X(89)90042-1

Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, *71*(356), 791–799. https://doi.org/10.1080/01621459.1976.10480949

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Boyd, R. (2017). A different kind of animal. In A Different Kind of Animal. Princeton University Press. https://doi.org/10.1515/9781400888528

Boyd, R., & Richerson, P. J. (1988). *Culture and the Evolutionary Process*. University of Chicago Press.

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., Adcock, R. A., Barch, D. M., Botvinick, M. M., Carver, C. S., Cools, R., Custers, R., Dickinson, A., Dweck, C. S., Fishbach, A., Gollwitzer, P. M., Hess, T. M., Isaacowitz, D. M., Mather, M., … for the MOMCAI group. (2014). Mechanisms of motivation–cognition interaction: Challenges and opportunities. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 443–472. https://doi.org/10.3758/s13415-014-0300-0

Breton, J. M., Eisner, J. S., Gandhi, V. S., Musick, N., Zhang, A., Long, K. L. P., Perloff, O. S., Hu, K. Y., Pham, C. M., Lalchandani, P., Barraza, M. K., Kantor, B., Kaufer, D., & Ben-Ami Bartal, I. (2022). Neural activation associated with outgroup helping in adolescent rats. *IScience*, *25*(6), 104412. https://doi.org/10.1016/j.isci.2022.104412

Briefer, E. F. (2018). Vocal contagion of emotions in non-human animals. *Proceedings of the Royal Society B: Biological Sciences, 285*(1873), 20172783. https://doi.org/10.1098/rspb.2017.2783

Brown, S. L., & Cialdini, R. B. (2015). Functional motives and functional consequences of prosocial behavior. In *The Oxford handbook of prosocial behavior* (pp. 346–361). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399813.001.0001

Buckner, R. L., & DiNicola, L. M. (2019). The brain's default network: Updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience*, *20*(10), 593–608. https://doi.org/10.1038/s41583-019-0212-7

Cacioppo, J. T. (2002). Social neuroscience: Understanding the pieces fosters understanding the whole and vice versa. *American Psychologist*, *57*, 819–831. https://doi.org/10.1037/0003-066X.57.11.819

Cacioppo, J. T., Amaral, D. G., Blanchard, J. J., Cameron, J. L., Carter, C. S., Crews, D., Fiske, S., Heatherton, T., Johnson, M. K., Kozak, M. J., Levenson, R. W., Lord, C., Miller, E. K., Ochsner, K., Raichle, M. E., Shea, M. T., Taylor, S. E., Young, L. J., & Quinn, K. J. (2007). Social neuroscience: Progress and implications for mental health. *Perspectives on Psychological Science*, *2*(2), 99–123. https://doi.org/10.1111/j.1745-6916.2007.00032.x

Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist*, *47*, 1019–1028. https://doi.org/10.1037/0003-066X.47.8.1019

Cacioppo, J. T., Berntson, G. G., & Decety, J. (2010). Social neuroscience and its relationship to social psychology. *Social Cognition*, *28*(6), 675–685. https://doi.org/10.1521/soco.2010.28.6.675

Cacioppo, J. T., & Decety, J. (2009). What are the brain mechanisms on which psychological processes are based? *Perspectives on Psychological Science, 4*(1), 10-18. https://www.jstor.org/stable/40212280

Cacioppo, J. T., & Decety, J. (2011). Social neuroscience: Challenges and opportunities in the study of complex behavior. *Annals of the New York Academy of Sciences*, *1224*(1), 162–173. https://doi.org/10.1111/j.1749-6632.2010.05858.x

Cacioppo, J. T., & Hawkley, L. C. (2009). Perceived social isolation and cognition. *Trends in Cognitive Sciences*, *13*(10), 447–454. https://doi.org/10.1016/j.tics.2009.06.005

Cacioppo, J. T., Norris, C. J., Decety, J., Monteleone, G., & Nusbaum, H. (2009). In the eye of the beholder: Individual differences in perceived social isolation predict regional brain activation to social stimuli. *Journal of Cognitive Neuroscience*, *21*(1), 83–92. https://doi.org/10.1162/jocn.2009.21007

Cacioppo, J. T., & Patrick, W. (2008). *Loneliness: Human Nature and the Need for Social Connection*. New York: Norton.

Calabro, R., Lyu, Y., & Leong, Y. C. (2022). Trial-by-trial fluctuations in amygdala activity track motivational enhancement of desirable sensory evidence during perceptual decision-making. *Cerebral Cortex*, bhac452. https://doi.org/10.1093/cercor/bhac452

Cao, Y., Contreras-Huerta, L. S., McFadyen, J., & Cunnington, R. (2015). Racial bias in neural response to others' pain is reduced with other-race contact. *Cortex, 70*, 68-78. https://doi.org/10.1016/j.cortex.2015.02.010

Caviola, L., & Faulmüller, N. (2014). Moral hypocrisy in economic games—How prosocial behavior is shaped by social expectations. *Frontiers in Psychology*, *5*, 897. https://doi.org/10.3389/fpsyg.2014.00897

Chang, L. J., Jolly, E., Cheong, J. H., Rapuano, K. M., Greenstein, N., Chen, P.-H. A., & Manning, J. R. (2021). Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *Science Advances*, *7*(17), 7129. https://doi.org/10.1126/sciadv.abf7129

Charpentier, C. J., & O'Doherty, J. P. (2018). The application of computational models to social neuroscience: Promises and pitfalls. *Social Neuroscience*, *13*(6), 637–647. https://doi.org/10.1080/17470919.2018.1518834

Chekroud, A., Everett, J., Bridge, H., & Hewstone, M. (2014). A review of neuroimaging studies of race-related prejudice: Does amygdala response reflect threat? *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00179

Chen, C., Decety, J., Huang, P. C., Chen, C. Y., & Cheng, Y. (2016). Testosterone administration in females modulates moral judgment and patterns of brain activation and functional connectivity. *Human Brain Mapping*, *37*, 3417–3430. https://doi.org/10.1002/hbm.23249

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Chen, P. H. A., Fareri, D., Güroğlu, B., Delgado, M. R., & Chang, L. J. (2021). Towards a Neurometric-based Construct Validity of Trust. *bioRxiv*. https://doi.org/10.1101/2021.07.04.451074

Cheney, D., Seyfarth, R., & Smuts, B. (1986). Social relationships and social cognition in nonhuman primates. *Science*, *234*(4782), 1361–1366. https://doi.org/10.1126/science.3538419

Cheong, J. H., Molani, Z., Sadhukha, S., & Chang, L. J. (2020). *Synchronized Affect in Shared Experiences Strengthens Social Connection*. PsyArXiv. https://doi.org/10.31234/osf.io/bd9wn

Cheong, J. H., Jolly, E., Sul, S., & Chang, L. J. (2017). Computational Models in Social Neuroscience. In *Computational Models of Brain and Behavior* (pp. 229–244). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119159193.ch17

Chiao, J. Y. (2010). Neural basis of social status hierarchy across species. *Current Opinion in Neurobiology*, *20*, 803–809. https://doi.org/10.1016/j.conb.2010.08.006

Chiao, J. Y., Iidaka, T., Gordon, H. L., Nogawa, J., Bar, M., Aminoff, E., Sadato, N., & Ambady, N. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience*, *20*(12), 2167–2174. https://doi.org/10.1162/jocn.2008.20151

Cikara, M., & van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science*, *9*(3), 245–274. https://doi.org/10.1177/1745691614527464

Cloutier, J., Li, T., & Correll, J. (2014). The impact of childhood experience on amygdala response to perceptually familiar black and white faces. *Journal of Cognitive Neuroscience*, *26*(9), 1992–2004. https://doi.org/10.1162/jocn_a_00605

Coan, J. A., Schaefer, H. S., & Davidson, R. J. (2006). Lending a hand: Social regulation of the neural response to threat. *Psychological Science*, *17*(12), 1032–1039. https://doi.org/10.1111/j.1467-9280.2006.01832.x

Cohen, S. (2004). Social relationships and health. *American Psychologist*, *59*(8), 676–684. https://doi.org/10.1177/0022146510383501

Cole, S. W., Hawkley, L. C., Arevalo, J. M., Sung, C. Y., Rose, R. M., & Cacioppo, J. T. (2007). Social regulation of gene expression in human leukocytes. *Genome Biology*, *8*(9), 1–13. https://doi.org/10.1186/gb-2007-8-9-r189

Contreras-Huerta, L. S., Baker, K. S., Reynolds, K. J., Batalha, L., & Cunnington, R. (2013). Racial bias in neural empathic responses to pain. *PLOS ONE*, *8*(12), 84001. https://doi.org/10.1371/journal.pone.0084001

Corradi-Dell'Acqua, C., Hofstetter, C., & Vuilleumier, P. (2011). Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. *Journal of Neuroscience*,

31(49), 17996–18006. https://doi.org/10.1523/JNEUROSCI.2686-11.2011

Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, *13*(1), 36–43. https://doi.org/10.1016/j.tics.2008.10.002

Cowell, J. M., Sommerville, J. A., & Decety, J. (2019). That's not fair: Children's neural computations of fairness and their impact on resource allocation behaviors and judgments. *Developmental Psychology*, *55*(11), 2299–2310. https://doi.org/10.1037/dev0000813

Crespi, B. J. (2001). The evolution of social behavior in microorganisms. *Trends in Ecology & Evolution*, *16*(4), 178–183. https://doi.org/10.1016/S0169-5347(01)02115-2

Crockett, M. J., Apergis-Schoute, A., Herrmann, B., Lieberman, M. D., Müller, U., & Robbins, T. W. (2013). Serotonin modulates striatal responses to fairness and retaliation in humans. *Journal of Neuroscience*, *33*, 3505–3513. https://doi.org/10.1523/JNEUROSCI.2761-12.2013

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences*, *107*(40), 17433-17438. https://doi.org/10.1073/pnas.100939610

Crockett, M. J., & Rini, R. A. (2015). Neuromodulators and the (in)stability of moral cognition. In J. Decety (Ed.), *The moral brain – A multidisciplinary perspective* (pp. 221–235). MIT Press.

Cross, K. P. (1977). Not can, but will college teaching be improved? *New Directions for Higher Education*, *1977*(17), 1–15. https://doi.org/10.1002/he.36919771703

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the bias map. *Advances in Experimental Social Psychology*, *40*, 61–149. https://doi.org/10.1016/S0065-2601(07)00002-0

Cundiff, J. M., & Matthews, K. A. (2017). Is subjective social status a unique correlate of physical health? A meta-analysis. *Health Psychology*, *36*(12), 1109–1125. https://doi.org/10.1037/hea0000534

Cunningham, W. A., & Brosch, T. (2012). Motivational salience: Amygdala tuning from traits, needs, values, and goals. *Current Directions in Psychological Science*, *21*(1), 54–59. https://doi.org/10.1177/0963721411430832

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, *15*(12), 806–813. https://doi.org/10.1111/j.0956-7976.2004.00760.x

Cutler, J., & Campbell-Meiklejohn, D. (2019). A comparative fMRI meta-analysis of altruistic and strategic decisions to give. *NeuroImage*, *184*, 227-241. https://doi.org/10.1016/j.neuroimage.2018.09.009

de Bruin, D., van Baar, J. M., Rodríguez, P. L., & FeldmanHall, O. (2023). Shared neural representations and temporal segmentation of political content predict ideological similarity. *Science Advances*, *9*(5), eabq5920. https://doi.org/10.1126/sciadv.abq5920

De Dreu, C. K., Greer, L. L., Van Kleef, G. A., Shalvi, S., & Handgraaf, M. J. (2011). Oxytocin promotes human ethnocentrism. *Proceedings of the National Academy of Sciences, 108*(4), 1262-1266. https://doi.org/10.1073/pnas.1015316108

De Dreu, C. K., Shalvi, S., Greer, L. L., Van Kleef, G. A., & Handgraaf, M. J. (2012). Oxytocin motivates non-cooperation in intergroup conflict to protect vulnerable in-group members. *PLOS ONE, 7*(11), e46751. https://doi.org/10.1371/journal.pone.0046751

Decety, J. (2015). The neural pathways, development and functions of empathy. *Current Opinion in Behavioral Sciences*, *3*, 1–6. https://doi.org/10.1016/j.cobeha.2014.12.001

Decety, J., Steinbeis, N., & Cowell, J. (2021). The neuro-development of social preferences in early childhood. *Current Opinion in Neurobiology, 68*, 23-28. https://doi.org/10.1016/j.conb.2020.12.009

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Decety, J., Norman, G. J., Berntson, G. G., & Cacioppo, J. T. (2012). A neurobehavioral evolutionary perspective on the mechanisms underlying empathy. *Progress in Neurobiology, 98*(1), 38-48. https://doi.org/10.1016/j.pneurobio.2012.05.001

Decety, J., & Porges, E. C. (2011). Imagining being the agent of actions that carry different moral consequences: an fMRI study. *Neuropsychologia, 49*(11), 2994-3001. https://doi.org/10.1016/j.neuropsychologia.2011.06.024

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3–6), 487–506. https://doi.org/10.1080/02643290244000239

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569–572. https://doi.org/10.1016/S0163-6383(97)90048-5

Dieffenbach, M. C., Gillespie, G. S. R., Burns, S. M., McCulloh, I. A., Ames, D. L., Dagher, M. M., Falk, E. B., & Lieberman, M. D. (2021). Neural reference groups: A synchrony-based classification approach for predicting attitudes using fNIRS. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 117–128. https://doi.org/10.1093/scan/nsaa115

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., ... & Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, 27(9), 1375-1380. https://doi.org/10.1016/j.cub.2017.04.002

Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research*, *1299*, 74–94. https://doi.org/10.1016/j.brainres.2009.07.007

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves "mind-reading" in humans. *Biological Psychiatry, 61*, 731–733.

https://doi.org/10.1016/j.biopsych.2006.07.015

Donaldson, Z. R., & Young, L. J. (2008). Oxytocin, vasopressin, and the neurogenetics of sociality. *Science, 322*(5903), 900-904. https://doi.org/10.1126/science.1158668

Dravida, S., Noah, J. A., Zhang, X., & Hirsch, J. (2020). Joint attention during live person-to-person contact activates rTPJ, including a sub-component associated with spontaneous eye-to-eye contact. *Frontiers in Human Neuroscience*, *14*. https://www.frontiersin.org/articles/10.3389/fnhum.2020.00201

Dreher, J. C., Dunne, S., Pazderska, A., Frodl, T., Nolan, J. J., & O'Doherty, J. P. (2016). Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males. *Proceedings of the National Academy of Sciences*, *113*(41), 11633–11638. https://doi.org/10.1073/pnas.160808511

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cognition*, *6*(5), 509–540. https://doi.org/10.1080/135062899394920

Du, E., & Chang, S. W. (2015). Neural components of altruistic punishment. *Frontiers in Neuroscience*, *9*, 26. https://doi.org/10.3389/fnins.2015.00026

Dunning, D. (2015). Motivated cognition in self and social thought. In *APA handbook of personality and social psychology, Volume 1: Attitudes and social cognition* (pp. 777–803). American Psychological Association. https://doi.org/10.1037/14341-025

Dunning, D., & Balcetis, E. (2013). Wishful seeing: How preferences shape visual perception. *Current Directions in Psychological Science*, *22*(1), 33–37. https://doi.org/10.1177/0963721412463693

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Dwortz, M. F., Curley, J. P., Tye, K. M., & Padilla-Coreano, N. (2022). Neural systems that facilitate the representation of social rank. *Philosophical Transactions of the Royal Society B*, *377*(1845), 20200444. https://doi.org/10.1098/rstb.2020.0444

Eckstein, M., Bamert, V., Stephens, S., Wallen, K., Young, L. J., Ehlert, U., & Ditzen, B. (2019). Oxytocin increases eye-gaze towards novel social and non-social stimuli. *Social Neuroscience, 14*(5), 594-607. https://doi.org/10.1080/17470919.2018.1542341

Edwards, D. H., & Kravitz, E. A. (1997). Serotonin, social status and aggression. *Current Opinion in Neurobiology*, *7*(6), 812–819. https://doi.org/10.1016/S0959-4388(97)80140-7

Erdelyi, M. H. (1974). A new look at the new look: Perceptual defense and vigilance. *Psychological Review*, *81*(1), 1–25. https://doi.org/10.1037/h0035852

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980–994.

Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. In S. C. Kolm & J. M. Ythier (Eds.), *Handbook of the economics of giving, altruism and reciprocity* (pp. 615–691). New York: Elsevier.

Feng, L., Liu, J., Wang, Z., Li, J., Li, L., Ge, L., Tian, J., & Lee, K. (2011). The other face of the other-race effect: An fMRI investigation of the other-race face categorization advantage. *Neuropsychologia*, *49*(13), 3739–3749. https://doi.org/10.1016/j.neuropsychologia.2011.09.031

Finn, E. S., & Bandettini, P. A. (2021). Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage*, *235*, 117963. https://doi.org/10.1016/j.neuroimage.2021.117963

Finn, E. S., Corlett, P. R., Chen, G., Bandettini, P. A., & Constable, R. T. (2018). Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04387-2

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X15000965

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1–74. https://doi.org/10.1016/S0065-2601(08)60317-2

Floresco, S. B. (2015). The nucleus accumbens: An interface between cognition, emotion, and action. *Annual Review of Psychology*, *66*(1), 25–52. https://doi.org/10.1146/annurev-psych-010213-115159

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, *67*(1), 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

Fourie, M. M., Subramoney, S., & Gobodo-Madikizela, P. (2017). A less attractive feature of empathy: Intergroup empathy bias. In M. Kondo (Ed.), Empathy: An Evidence-based Interdisciplinary Perspective (pp. 45-61): InTech. ISBN 978-953-51-5420-4.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological review*, 118(2), 247. https://doi.org/10.1037/a0022327

Freeman, J. B., Stolier, R. M., Brooks, J. A., & Stillerman, B. S. (2018). The neural representational geometry of social perception. *Current Opinion in Psychology*, *24*, 83-91. https://doi.org/10.1016/j.copsyc.2018.10.003

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current Biology*, *17*(16), 724–732. https://doi.org/10.1016/j.cub.2007.05.068

Genevsky, A., Västfjäll, D., Slovic, P., & Knutson, B. (2013). Neural underpinnings of the identifiable victim effect: Affect shifts preferences for giving. *Journal of Neuroscience*, *33*(43), 17188–

17196. https://doi.org/10.1523/JNEUROSCI.2348-13.2013

Gintis, H. (2011). Gene–culture coevolution and the nature of human sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 878–888. https://doi.org/10.1098/rstb.2010.0310

Glänzel, W., & Schubert, A. (2005). Analysing scientific networks through co-authorship. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems* (pp. 257–276). Springer Netherlands. https://doi.org/10.1007/1-4020-2755-9_12

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(supplement_3), 15647–15654. https://doi.org/10.1073/pnas.1014269108

Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience*, *4*(8), 845–850. https://doi.org/10.1038/90565

Gottlieb, J., Lopes, M., & Oudeyer, P.-Y. (2016). Motivated cognition: Neural and computational mechanisms of curiosity, attention, and intrinsic motivation. In *Recent Developments in Neuroscience Research on Human Motivation* (Vol. 19, pp. 149–172). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-742320160000019017

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*(2), 159–166. https://doi.org/10.1016/S0959-4388(03)00040-0

Gritsun, T. A., Le Feber, J., Stegenga, J., & Rutten, W. L. C. (2010). Network bursts in cortical cultures are best simulated using pacemaker neurons and adaptive synapses. *Biological Cybernetics*, *102*(4), 293–310. https://doi.org/10.1007/s00422-010-0366-x

Hackel, L. M., & Amodio, D. M. (2018). Computational neuroscience approaches to social cognition. *Current Opinion in Psychology*, *24*, 92–97. https://doi.org/10.1016/j.copsyc.2018.09.001

Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: Dissociable neural correlates and effects on choice. *Nature Neuroscience*, *18*(9), Article 9. https://doi.org/10.1038/nn.4080

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659. https://doi.org/10.1038/nrn2897

Hakonen, M., Ikäheimonen, A., Hultèn, A., Kauttonen, J., Koskinen, M., Lin, F.-H., Lowe, A., Sams, M., & Jääskeläinen, I. P. (2022). Processing of an audiobook in the human brain is shaped by cultural family background. *Brain Sciences*, *12*(5), 649. https://doi.org/10.3390/brainsci12050649

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport*,

11(11), 2351–2354. https://links.lww.com/WNR/A4

Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *Journal of Neuroscience*, *34*(3), 698–704. https://doi.org/10.1523/JNEUROSCI.2489-13.2014

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Haufe, S., DeGuzman, P., Henin, S., Arcaro, M., Honey, C. J., Hasson, U., & Parra, L. C. (2018). Elucidating relations between fMRI, ECoG, and EEG through a common natural stimulus. *NeuroImage*, *179*, 79–91. https://doi.org/10.1016/j.neuroimage.2018.06.016

Hawley, P. H. (2015). Social dominance in childhood and its evolutionary underpinnings: Why it matters and what we can do. *Pediatrics*, *135*, 31–38. https://doi.org/10.1542/peds.2014-3549D

Heider, F. (1946). Attitudes and Cognitive Organization. *The Journal of Psychology*, *21*(1), 107–112. https://doi.org/10.1080/00223980.1946.9917275

Hein, G., Engelmann, J. B., Vollberg, M. C., & Tobler, P. N. (2016). How learning shapes the empathic brain. *Proceedings of the National Academy of Sciences*, *113*(1), 80–85. https://doi.org/10.1073/pnas.1514539112

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149–160. https://doi.org/10.1016/j.neuron.2010.09.003

Henrich, J. (2016). The secret of our success. Princeton: Princeton University Press.

Higgins, E. T., & Molden, D. C. (2003). How strategies for making judgments and decisions affect cognition: Motivated cognition revisited. In *Foundations of Social Cognition: A Festschrift in Honor of Robert S. Wyer, Jr* (pp. 211–235). New York: Lawrence Erlbaum Associates Publishers.

Hirsch, J., Zhang, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *NeuroImage*, *157*, 314–330. https://doi.org/10.1016/j.neuroimage.2017.06.018

Holt-Lunstad, J. (2021). Loneliness and social isolation as risk factors: The power of social connection in prevention. *American Journal of Lifestyle Medicine*, *15*(5), 567–573. https://doi.org/10.1177/15598276211009454

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science*, *10*(2), 227–237. https://doi.org/10.1177/1745691614568352

Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *Journal of Neuroscience*, *32*(44), 15277–15283.

https://doi.org/10.1523/JNEUROSCI.1800-12.2012

Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A Neurocomputational Model of Altruistic Choice and Its Implications. *Neuron*, *87*(2), 451–462. https://doi.org/10.1016/j.neuron.2015.06.031

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187. https://doi.org/10.1037/a0020463

Hughes, B. L., & Beer, J. S. (2012). Orbitofrontal cortex and anterior cingulate cortex are modulated by motivated social cognition. *Cerebral Cortex*, *22*(6), 1372–1381. https://doi.org/10.1093/cercor/bhr213

Hughes, B. L., Camp, N. P., Gomez, J., Natu, V. S., Grill-Spector, K., & Eberhardt, J. L. (2019). Neural adaptation to faces reveals racial outgroup homogeneity effects in early perception. *Proceedings of the National Academy of Sciences*, *116*(29), 14532–14537. https://doi.org/10.1073/pnas.1822084116

Hughes, B. L., & Zaki, J. (2015). The neuroscience of motivated cognition. *Trends in Cognitive Sciences*, *19*(2), 62–64. https://doi.org/10.1016/j.tics.2014.12.006

Huppert, E., Cowell, J. M., Cheng, Y., Contreras-Ibáñez, C., Gomez-Sicard, N., Gonzalez-Gaeda, M. L., Huepe, D., Ibanez, A., Lee, K., Mahasneh, R., Malcolm-Smith, S., Salas, N., Selcuk, B., Tungodden, B., Wong, A., Zhou, X., & Decety, J. (2019). The development of children's preferences for equality and equity across 13 individualistic and collectivist cultures. *Developmental Science, 22*(2) e1279. https://doi.org/10.1111/desc.12729

Hutzler, F. (2014). Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *NeuroImage*, *84*, 1061–1069. https://doi.org/10.1016/j.neuroimage.2012.12.075

Hyon, R., Kleinbaum, A. M., & Parkinson, C. (2020). Social network proximity predicts similar trajectories of psychological states: Evidence from multi-voxel spatiotemporal dynamics. *NeuroImage*, *216*, 116492. https://doi.org/10.1016/j.neuroimage.2019.116492

Israel, S., Weisel, O., Ebstein, R. P., & Bornstein, G. (2012). Oxytocin, but not vasopressin, increases both parochial and universal altruism. *Psychoneuroendocrinology, 37*(8), 1341-1344. https://doi.org/10.1016/j.psyneuen.2012.02.001

Izuma, K. (2013). The neural basis of social influence and attitude change. *Current Opinion in Neurobiology*, *23*(3), 456–462. https://doi.org/10.1016/j.conb.2013.03.009

Izuma, K. (2017). Chapter 16—The neural bases of social influence on valuation and behavior. In J.-C. Dreher & L. Tremblay (Eds.), *Decision Neuroscience* (pp. 199–209). New York: Academic Press. https://doi.org/10.1016/B978-0-12-805308-9.00016-6

Izuma, K., & Adolphs, R. (2013). Social manipulation of preference in the human brain. *Neuron*, *78*(3), 563–573. https://doi.org/10.1016/j.neuron.2013.03.023

Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., & Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences*, *107*(51), 22014–22019. https://doi.org/10.1073/pnas.1011879108

Izuma, K., Saito, D. N., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience*, *22*(4), 621–631.

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, *24*(3), 771–779. https://doi.org/10.1016/j.neuroimage.2004.09.006

Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *WIREs Cognitive Science*, *2*(1), 8–21. https://doi.org/10.1002/wcs.80

Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., & Lu, C. (2015). Leader emergence through interpersonal neural synchronization. *Proceedings of the National Academy of Sciences*, *112*(14), 4274–4279. https://doi.org/10.1073/pnas.1422930112

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. *Journal of Neuroscience*, *32*(45), 16064–16069. https://doi.org/10.1523/JNEUROSCI.2926-12.2012

Jiang, Y., Mi, Q., & Zhu, L. (2023). Neurocomputational mechanism of real-time distributed learning on social networks. *Nature Neuroscience*, *26*(3), Article 3. https://doi.org/10.1038/s41593-023-01258-y

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General*, *146*(4), 569–576. https://doi.org/10.1037/xge0000271

Kanitz, E., Tuchscherer, M., Puppe, B., Tuchscherer, A., & Stabenow, B. (2004). Consequences of repeated early isolation in domestic piglets on their behavioural, neuroendocrine, and immunological responses. *Brain, Behavior, and Immunity*, *18*(1), 35-45. https://doi.org/10.1016/s0889-1591(03)00085-0

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and Theory of Mind. *NeuroImage*, *122*, 6-19. https://doi.org/10.1016/j.neuroimage.2015.07.082

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Karelina, K., Norman, G. J., Zhang, N., Morris, J. S., Peng, H., & DeVries, A. C. (2009). Social isolation alters neuroinflammatory response to stroke. *Proceedings of the National Academy of Sciences, 106*(14), 5895-5900. https://doi.org/10.1073/pnas.081073710

Katabi, N., Simon, H., Yakim, S., Ravreby, I., Ohad, T., & Yeshurun, Y. (2023). Deeper Than You Think: Partisanship-Dependent Brain Responses in Early Sensory and Motor Brain Regions. *Journal of Neuroscience*, *43*(6), 1027–1037. https://doi.org/10.1523/JNEUROSCI.0895-22.2022

Kinreich, S., Djalovski, A., Kraus, L., Louzoun, Y., & Feldman, R. (2017). Brain-to-brain synchrony during naturalistic social interactions. *Scientific Reports*, *7*(1), 17060. https://doi.org/10.1038/s41598-017-17339-5

Klein, J. T., & Platt, M. L. (2013). Social information signaling by neurons in primate striatum. *Current Biology*, *23*(8), 691–696. https://doi.org/10.1016/j.cub.2013.03.022

Kljajevic, V. (2021). Neurology of humor. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 5387–5394). Springer International Publishing. https://doi.org/10.1007/978-3-319-19650-3_3242

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., & Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, *61*(1), 140–151. https://doi.org/10.1016/j.neuron.2008.11.027

Klucharev, V., Munneke, M. A. M., Smidts, A., & Fernández, G. (2011). Downregulation of the posterior medial frontal cortex prevents social conformity. *Journal of Neuroscience*, *31*(33), 11934–11940. https://doi.org/10.1523/JNEUROSCI.1869-11.2011

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8514–8519. https://doi.org/10.1073/pnas.10169551

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435*(7042), 673-676. https://doi.org/10.1038/nature03701

Koski, J. E., Collins, J. A., & Olson, I. R. (2017). The neural representation of social status in the extended face-processing network. *European Journal of Neuroscience*, *46*(12), 2795–2806. https://doi.org/10.1111/ejn.13770

Koski, J. E., Xie, H., & Olson, I. R. (2015). Understanding social hierarchies: The neural and psychological foundations of status perception. *Social Neuroscience*, *10*(5), 527–550. https://doi.org/10.1080/17470919.2015.1013223

Krahé, C., Springer, A., Weinman, J. A., & Fotopoulou, A. (2013). The social modulation of pain: Others as predictive signals of salience–A systematic review. *Frontiers in Human Neuroscience*, *7*, 386. https://doi.org/10.3389/fnhum.2013.00386

Krolak-Salmon, P., Hénaff, M. A., Vighetto, A., Bauchet, F., Bertrand, O., Mauguière, F., & Isnard, J. (2006). Experiencing and detecting happiness in humans: The role of the supplementary

motor area. *Annals of Neurology*, *59*(1), 196–199. https://doi.org/10.1002/ana.20706

Kruglanski, A. W. (1999). Motivation, cognition, and reality: Three memos for the next generation of research. *Psychological Inquiry*, *10*(1), 54–58. https://doi.org/10.1207/s15327965pli1001_8

Kurtz, L. E., & Algoe, S. B. (2015). Putting laughter in context: Shared laughter as behavioral indicator of relationship well-being. *Personal Relationships*, *22*(4), 573–590. https://doi.org/10.1111/pere.12095

Kweon, H., Aydogan, G., Dagher, A., Bzdok, D., Ruff, C. C., Nave, G., & Koellinger, P. D. (2022). Human brain anatomy reflects separable genetic and environmental components of socioeconomic status. *Science Advances*, *8*(20), 2923. https://doi.org/10.1126/sciadv.abm2923

Lahnakoski, J. M., Glerean, E., Jääskeläinen, I. P., Hyönä, J., Hari, R., Sams, M., & Nummenmaa, L. (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage*, *100*, 316–324. https://doi.org/10.1016/j.neuroimage.2014.06.022

Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: Bringing genetics and the human sciences together. *Nature Reviews Genetics*, *11*(2), 137–148. https://doi.org/10.1038/nrg2734

Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience, 19*(1), 42-58. https://doi.org/10.1162/jocn.2007.19.1.42

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*, 541–567. https://doi.org/10.1080/135062899394939

Lau, T., Pouncy, H. T., Gershman, S. J., & Cikara, M. (2018). Discovering social groups via latent structure learning. *Journal of Experimental Psychology: General*, *147*, 1881–1891. https://doi.org/10.1037/xge0000470

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*(1), 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

Lee, Y. A., & Goto, Y. (2018). The roles of serotonin in decision-making under social group conditions. *Scientific Reports*, *8*, 10704. https://doi.org/10.1038/s41598-018-29055-9

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, *1*(4), 1–9. https://doi.org/10.1038/s41562-017-0067

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences*, *114*(50), 13290–13295. https://doi.org/10.1073/pnas.1702493114

Leong, Y. C., Chen, J., Willer, R., & Zaki, J. (2020). Conservative and liberal attitudes drive polarized neural responses to political content. *Proceedings of the National Academy of Sciences*, *117*(44), 27731–27739. https://doi.org/10.1073/pnas.2008530117

Leong, Y. C., Dziembaj, R., & D'Esposito, M. (2021). Pupil-Linked Arousal Biases Evidence Accumulation Toward Desirable Percepts During Perceptual Decision-Making. *Psychological Science*, *32*(9), 1494–1509. https://doi.org/10.1177/09567976211004547

Leong, Y. C., Hughes, B. L., Wang, Y., & Zaki, J. (2019). Neurocomputational mechanisms underlying motivated seeing. *Nature Human Behaviour*, *3*(9), 962–973. https://doi.org/10.1038/s41562-019-0637-z

Leong, Y. C., & Zaki, J. (2018). Unrealistic optimism in advice taking: A computational account. *Journal of Experimental Psychology: General*, *147*(2), 170–189. https://doi.org/10.1037/xge0000382

Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, *8*(6), 720–722. https://doi.org/10.1038/nn1465

Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is There a 'Social' Brain? Implementations and Algorithms. *Trends in Cognitive Sciences*, *24*(10), 802–813. https://doi.org/10.1016/j.tics.2020.06.011

Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences*, *113*(35), 9763–9768. https://doi.org/10.1073/pnas.1603198113

MacDougall-Shackleton, S. A. (2011). The levels of analysis revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1574), 2076–2085. https://doi.org/10.1098/rstb.2010.0363

Madan, C. R. (2017). Motivated cognition: Effects of reward, emotion, and other motivational factors across a variety of cognitive domains. *Collabra: Psychology*, *3*(1), 24. https://doi.org/10.1525/collabra.111

Marder, E., Goeritz, M. L., & Otopalik, A. G. (2015). Robust circuit rhythms in small circuits arise from variable circuit components and mechanisms. *Current Opinion in Neurobiology*, *31*, 156–163. https://doi.org/10.1016/j.conb.2014.10.012

Mattan, B. D., Kubota, J. T., & Cloutier, J. (2017). How social status shapes person perception and evaluation: A social neuroscience perspective. *Perspectives on Psychological Science*, *12*(3), 468–507. https://doi.org/10.1177/1745691616677828

Mazur, A., Booth, A., & Dabbs Jr, J. M. (1992). Testosterone and chess competition. *Social Psychology Quarterly*, 70–77. https://doi.org/10.3390/ijerph17041204

McCormick, I. A., Walkey, F. H., & Green, D. E. (1986). Comparative perceptions of driver ability—A confirmation and expansion. *Accident Analysis & Prevention*, *18*(3), 205–208.

https://doi.org/10.1016/0001-4575(86)90004-7

McEwen, B. S. (2012). Brain on stress: How the social environment gets under the skin. *Proceedings of the National Academy of Sciences*, *109*, 17180–17185. https://doi.org/10.1073/pnas.1121254109

McEwen, B. S., Bowles, N. P., Gray, J. D., Hill, M. N., Hunter, R. G., Karatsoreos, I. N., & Nasca, C. (2015). Mechanisms of stress in the brain. *Nature Neuroscience*, *18*(10), 1353–1363. https://doi.org/10.1038/nn.4086

Meyer, M. L. (2019). Social by default: Characterizing the social functions of the resting brain. *Current Directions in Psychological Science*, *28*(4), 380–386. https://doi.org/10.1177/0963721419857759

Michalska, K. J., Decety, J., Liu, C., Chen, Q., Martz, M. E., Jacob, S., Hipwell, A. E., Lee, S. S., Chronis-Tuscano, A., Waldman, I. D., & Lahey, B. B. (2014). Genetic imaging of the association of oxytocin receptor gene (OXTR) polymorphisms with positive maternal parenting. *Frontiers in Behavioral Neuroscience, 8*:21. https://doi.org/10.3389/fnbeh.2014.00021

Mikolajczak, M., Pinon, N., Lane, A., de Timary, P., & Luminet, O. (2010). Oxytocin not only increases trust when money is at stake, but also when confidential information is in the balance. *Biological Psychology, 85*(1), 182-184. https://doi.org/10.1016/j.biopsycho.2010.05.010

Molenberghs, P., & Louis, W. R. (2018). Insights from fMRI studies into ingroup bias. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01868

Moore, C., Dunham, P. J., & Dunham, P. (Eds.). (2014). *Joint Attention: Its Origins and Role in Development*. Psychology Press. https://doi.org/10.4324/9781315806617

Morelli, S. A., Leong, Y. C., Carlson, R. W., Kullar, M., & Zaki, J. (2018). Neural detection of socially valued community members. *Proceedings of the National Academy of Sciences*, *115*(32), 8149–8154. https://doi.org/10.1073/pnas.1712811115

Morelli, S. A., Ong, D. C., Makati, R., Jackson, M. O., & Zaki, J. (2017). Empathy and well-being correlate with centrality in different social networks. *Proceedings of the National Academy of Sciences*, *114*(37), 9843–9847. https://doi.org/10.1073/pnas.1702155114

Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annual Review of Neuroscience*, *31*(1), 69–89. https://doi.org/10.1146/annurev.neuro.31.061307.090723

Moutsiana, C., Charpentier, C. J., Garrett, N., Cohen, M. X., & Sharot, T. (2015). Human frontal–subcortical circuit and asymmetric belief updating. *Journal of Neuroscience*, *35*(42), 14077–14085. https://doi.org/10.1523/JNEUROSCI.1120-15.2015

Müller, V., Perdikis, D., Mende, M. A., & Lindenberger, U. (2021). Interacting brains coming in sync through their minds: An interbrain neurofeedback study. *Annals of the New York Academy of Sciences*, *1500*(1), 48–68. https://doi.org/10.1111/nyas.14605

Nakahashi, W., & Ohtsuki, H. (2015). When is emotional contagion adaptive? *Journal of Theoretical Biology*, *380*, 480–488. https://doi.org/10.1016/j.jtbi.2015.06.014

Nastase, S. A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, *14*(6), 667–685. https://doi.org/10.1093/scan/nsz037

Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A.-L. (2000). The sound of many hands clapping. *Nature*, *403*(6772), 849–850. https://doi.org/10.1038/35002660

Nelissen, R. M., & Meijers, M. H. (2011). Social benefits of luxury brands as costly signals of wealth and status. *Evolution and Human Behavior*, *32*(5), 343–355. https://doi.org/10.1016/j.evolhumbehav.2010.12.002

Nelson-Coffey, S. K., Fritz, M. M., Lyubomirsky, S., & Cole, S. W. (2017). Kindness in the blood: A randomized controlled trial of the gene regulatory impact of prosocial behavior. *Psychoneuroendocrinology*, *81*, 8–13. https://doi.org/10.1016/j.psyneuen.2017.03.025

Nessler, J. A., & Gilliland, S. J. (2009). Interpersonal synchronization during side by side treadmill walking is influenced by leg length differential and altered sensory feedback. *Human Movement Science*, *28*(6), 772–785. https://doi.org/10.1016/j.humov.2009.04.007

Nguyen, M., Chang, A., Micciche, E., Meshulam, M., Nastase, S. A., & Hasson, U. (2022). Teacher–student neural coupling during teaching and learning. *Social Cognitive and Affective Neuroscience*, *17*(4), 367–376. https://doi.org/10.1093/scan/nsab103

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, *184*, 161–170. https://doi.org/10.1016/j.neuroimage.2018.09.010

Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2021). Neural synchrony in mother–child conversation: Exploring the role of conversation patterns. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 93–102. https://doi.org/10.1093/scan/nsaa079

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Nonogaki, K., Nozue, K., & Oka, Y. (2007). Social isolation affects the development of obesity and type 2 diabetes in mice. *Endocrinology*, *148*(10), 4658-4666. https://doi.org/10.1210/en.2007-0296

Norman, G. J., Hawkley, L. C., Cole, S. W., Berntson, G. G., & Cacioppo, J. T. (2012). Social neuroscience: The social brain, oxytocin, and health. *Social Neuroscience*, *7*(1), 18–29. https://doi.org/10.1080/17470919.2011.568702

Novembre, G., & Iannetti, G. D. (2021). Hyperscanning alone cannot prove causality. Multibrain stimulation can. *Trends in Cognitive Sciences*, *25*(2), 96–99. https://doi.org/10.1016/j.tics.2020.11.003

Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I. P., Hari, R., & Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *Proceedings of the National Academy of Sciences*, *109*(24), 9599–9604. https://doi.org/10.1073/pnas.1206095109

Nummenmaa, L., Lahnakoski, J. M., & Glerean, E. (2018). Sharing the social world via intersubject neural synchronisation. *Current Opinion in Psychology*, *24*, 7–14. https://doi.org/10.1016/j.copsyc.2018.02.021

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249. https://doi.org/10.1016/j.tics.2005.03.010

Oliveira, T., Gouveia, M. J., & Oliveira, R. F. (2009). Testosterone responsiveness to winning and losing experiences in female soccer players. *Psychoneuroendocrinology*, *34*(7), 1056–1064. https://doi.org/10.1016/j.psyneuen.2009.02.006

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-Based fMRI and Its Application to Reward Learning and Decision Making. *Annals of the New York Academy of Sciences*, *1104*(1), 35–53. https://doi.org/10.1196/annals.1390.022

O'Reilly, R. C. (2020). Unraveling the mysteries of motivation. *Trends in Cognitive Sciences*, *24*(6), 425–434. https://doi.org/10.1016/j.tics.2020.03.001

Overgaauw, S., Güroğlu, B., & Crone, E. A. (2012). Fairness considerations when I know more than you do: Developmental comparisons. *Frontiers in Psychology*, *3*, 424. https://doi.org/10.3389/fpsyg.2012.00424

Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature Neuroscience*, *24*(9), 1292–1301. https://doi.org/10.1038/s41593-021-00916-3

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2017). Spontaneous neural encoding of social network position. *Nature Human Behaviour*, *1*(5), 1–7. https://doi.org/10.1038/s41562-017-0072

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2018). Similar neural responses predict friendship. *Nature Communications*, *9*(1), 332. https://doi.org/10.1038/s41467-017-02722-7

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a "low road" to "many roads" of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–782. https://doi.org/10.1038/nrn2920

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*(5), 729–738. https://doi.org/10.1162/089892900562552

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2020). Infant and adult brains are coupled to the dynamics of natural communication. *Psychological Science*, *31*(1), 6–17.

https://doi.org/10.1177/0956797619878698

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. https://doi.org/10.1016/j.tics.2005.12.004

Poletto, R., Steibel, J. P., Siegford, J. M., & Zanella, A. J. (2006). Effects of early weaning and social isolation on the expression of glucocorticoid and mineralocorticoid receptor and 11β-hydroxysteroid dehydrogenase 1 and 2 mRNAs in the frontal cortex and hippocampus of piglets. *Brain Research*, *1067*(1), 36-42. https://doi.org/10.1016/j.brainres.2005.10.001

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, *92*(3), 492–512. https://doi.org/10.1016/j.biopsycho.2012.02.007

Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., & Buck, A. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254–1258. https://doi.org/10.1126/science.1100735

Quintana, D. S., & Guastella, A. J. (2020). An allostatic theory of oxytocin. *Trends in Cognitive Sciences*, *24*(7), 515-528. https://doi.org/10.1016/j.tics.2020.03.008

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Richeson, J. A., Todd, A. R., Trawalter, S., & Baird, A. A. (2008). Eye-gaze direction modulates race-related amygdala activity. *Group Processes & Intergroup Relations*, *11*(2), 233–246. https://doi.org/10.1177/1368430207088040

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Thompson, R., Ditzen, B., Patel, R., & Pagnoni, G. (2012). Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology*, *37*(4), 447-461. https://doi.org/10.1016/j.psyneuen.2011.07.013

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405. https://doi.org/10.1016/S0896-6273(02)00755-9

Ronquillo, J., Denson, T. F., Lickel, B., Lu, Z.-L., Nandy, A., & Maddox, K. B. (2007). The effects of skin tone on race-related amygdala activity: An fMRI investigation. *Social Cognitive and Affective Neuroscience*, *2*(1), 39–44. https://doi.org/10.1093/scan/nsl043

Rose, M. C., Styr, B., Schmid, T. A., Elie, J. E., & Yartsev, M. M. (2021). Cortical representation of group social communication in bats. *Science*, *374*(6566), eaba9584. https://doi.org/10.1126/science.aba9584

Sachs, M. E., Habibi, A., Damasio, A., & Kaplan, J. T. (2020). Dynamic intersubject neural synchronization reflects affective responses to sad music. *NeuroImage*, *218*, 116512. https://doi.org/10.1016/j.neuroimage.2019.116512

Saito, T., Motoki, K., Nouchi, R., Kawashima, R., & Sugiura, M. (2020). Loneliness modulates automatic attention to warm and competent faces: Preliminary evidence from an eye-tracking study. *Frontiers in Psychology*, *10*, 2967. https://doi.org/10.3389/fpsyg.2019.02967

Salgado, S., & Berntsen, D. (2021). It won't happen to us": Unrealistic optimism affects COVID-19 risk assessments and attitudes regarding protective behaviour. *Journal of Applied Research in Memory and Cognition*, *10*(3), 368–380. https://doi.org/10.1016/j.jarmac.2021.07.006

Santos, A., Mier, D., Kirsch, P., & Meyer-Lindenberg, A. (2011). Evidence for a general face salience signal in human amygdala. *NeuroImage*, *54*(4), 3111–3116. https://doi.org/10.1016/j.neuroimage.2010.11.024

Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science*, *308*(5722), 648–652. https://doi.org/10.1126/science.110647s

Schafer, M., Kamilar-Britt, P., Sahani, V., Bachi, K., & Schiller, D. (2022). Hippocampal Place-like Signal in Latent Space (p. 2022 07 15 499827). biorxiv. https://doi.org/10.1101/2022.07.15.499827

Schafer, M., & Schiller, D. (2018). Navigating social space. *Neuron*, *100*(2), 476–489. https://doi.org/10.1016/j.neuron.2018.10.006

Schie, C. C., Chiu, C.-D., Rombouts, S. A. R. B., Heiser, W. J., & Elzinga, B. M. (2018). When compliments do not hit but critiques do: An fMRI study into self-esteem and self-knowledge in processing social feedback. *Social Cognitive and Affective Neuroscience*, *13*(4), 404–417. https://doi.org/10.1093/scan/nsy014

Schippers, M. B., Roebroeck, A., Renken, R., Nanetti, L., & Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences*, *107*(20), 9388–9393. https://doi.org/10.1073/pnas.1001791107

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Shamay-Tsoory, S. G., Abu-Akel, A., Palgi, S., Sulieman, R., Fischer-Shofty, M., Levkwitz, Y., & Decety, J. (2013). Giving peace a chance: Oxytocin increases empathy to pain in the context of the Israeli-Palestinian conflict. *Psychoneuroendocrinology*, *38*(12), 3139-3144. https://doi.org/10.1016/j.psyneuen.2013.09.015

Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R., & Dolan, R. J. (2012). How dopamine enhances an optimism bias in humans. *Current Biology*, *22*(16), 1477–1481. https://doi.org/10.1016/j.cub.2012.05.053

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, *109*(42), 17058–17062. https://doi.org/10.1073/pnas.1205828109

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. https://doi.org/10.1038/nn.2949

Shepperd, J. A., Klein, W. M. P., Waters, E. A., & Weinstein, N. D. (2013). Taking stock of unrealistic optimism. *Perspectives on Psychological Science*, *8*(4), 395–411. https://doi.org/10.1177/1745691613485247

Shepperd, J. A., Waters, E. A., Weinstein, N. D., & Klein, W. M. P. (2015). A primer on unrealistic optimism. *Current Directions in Psychological Science*, *24*(3), 232–237. https://doi.org/10.1177/0963721414568341

Sherman, M. T., Seth, A. K., & Kanai, R. (2016). Predictions shape confidence in right inferior frontal gyrus. *Journal of Neuroscience*, *36*(40), 10323–10336. https://doi.org/10.1523/JNEUROSCI.1092-16.2016

Shonkoff, J. P. (2003). From neurons to neighborhoods: Old and new challenges for developmental and behavioral pediatrics. *Journal of Developmental & Behavioral Pediatrics*, *24*(1), 70–76. https://doi.org/10.1097/00004703-200302000-00014

Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social bonds of female baboons enhance infant survival. *Science*, *302*(5648), 1231–1234. https://doi.org/ 10.1126/science.10885

Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., & Cheney, D. L. (2010). Strong and consistent social bonds enhance the longevity of female baboons. *Current Biology*, *20*(15), 1359–1361. https://doi.org/10.1016/j.cub.2010.05.06s

Simpson, B., & Willer, R. (2015). Beyond altruism: Sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology*, *41*(1), 43–63. https://doi.org/10.1146/annurev-soc-073014-112242

Somerville, L. H., Kelley, W. M., & Heatherton, T. F. (2010). Self-esteem modulates medial prefrontal cortical responses to evaluative social feedback. *Cerebral Cortex*, *20*(12), 3005–3013. https://doi.org/10.1093/cercor/bhq049

Son, J.-Y., Bhandari, A., & FeldmanHall, O. (2021). Cognitive maps of social features enable flexible inference in social networks. *Proceedings of the National Academy of Sciences*, *118*(39), e2021699118. https://doi.org/10.1073/pnas.2021699118

Song, C., Sandberg, K., Rutiku, R., & Kanai, R. (2022). Linking human behaviour to brain structure: Further challenges and possible solutions. *Nature Reviews Neuroscience*, *23*, 517–518. https://doi.org/10.1038/s41583-022-00614-4

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences*, *118*(33), 2021905118. https://doi.org/10.1073/pnas.2021905118

Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007). The neural signature of social norm compliance. *Neuron*, *56*(1), 185–196. https://doi.org/10.1016/j.neuron.2007.09.011

Stallen, M., & Sanfey, A. G. (2015). The neuroscience of social conformity: Implications for fundamental and applied research. *Frontiers in Neuroscience*, *9*. https://doi.org/10.3389/fnins.2015.00337

Staudinger, M. R., & Büchel, C. (2013). How initial confirmatory experience potentiates the detrimental influence of bad advice. *NeuroImage*, *76*, 125–133. https://doi.org/10.1016/j.neuroimage.2013.02.074

Stavans, M., & Baillargeon, R. (2019). Infants expect leaders to right wrongs. *Proceedings of the National Academy of Sciences, 116*(33), 16292-16301. https://doi.org/10.1073/pnas.182009111

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, *107*(32), 14425–14430. https://doi.org/10.1073/pnas.1008662107

Strobel, A., Zimmermann, J., Schmitz, A., Reuter, M., Lis, S., Windmann, S., & Kirsch, P. (2011). Beyond revenge: Neural and genetic bases of altruistic punishment. *Neuroimage*, *54*(1), 671–680. https://doi.org/10.1016/j.neuroimage.2010.07.051

Sun, P., Zheng, L., Li, L., Guo, X., Zhang, W., & Zheng, Y. (2016). The neural responses to social cooperation in gain and loss context. *PLoS One*, *11*(8), 0160503. https://doi.org/10.1371/journal.pone.0160503

Szymanska, M., Schneider, M., Chateau-Smith, C., Nezelof, S., & Vulliez-Coady, L. (2017). Psychophysiological effects of oxytocin on parent–child interactions: A literature review on oxytocin and parent–child interactions. *Psychiatry and Clinical Neurosciences, 71*(10), 690-705. https://doi.org/10.1111/pcn.12544

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A map for social navigation in the human brain. *Neuron*, *87*(1), 231–243. https://doi.org/10.1016/j.neuron.2015.06.011

Thomsen, L. (2020). The developmental origins of social hierarchy: How infants and young children mentally represent and respond to power and status. *Current Opinion in Psychology*, *33*, 201–208. https://doi.org/10.1016/j.copsyc.2019.07.044

Tin, F., Buyukoyer Dawkins, M., Stavans, M., & Baillargeon. (2020). Principles and concepts in early moral cognition. In J. Decety (Ed.), *The social brain – A developmental perspective. Cambridge (pp.41-65)*. Cambridge, MA: MIT Press.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Tomasello, M. (2020). The adaptive origins of uniquely human sociality. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190493. http://doi.org/10.1098/rstb.2019.0493

Tomashin, A., Gordon, I., & Wallot, S. (2022). Interpersonal physiological synchrony predicts group cohesion. *Frontiers in Human Neuroscience*, *16*, 439. https://doi.org/10.3389/fnhum.2022.903407

Tooley, U. A., Bassett, D. S., & Mackey, A. P. (2021). Environmental influences on the pace of brain development. *Nature Reviews Neuroscience*, *22*(6), 372–384. https://doi.org/10.1038/s41583-

021-00457-5

Tricomi, E., & Sullivan-Toole, H. (2015). Fairness and inequity aversion. *Brain Mapping: An Encyclopedic Reference*, *3*, 3–8. https://doi.org/10.1016/B978-0-12-397025-1.00142-1

Trost, W., Frühholz, S., Cochrane, T., Cojan, Y., & Vuilleumier, P. (2015). Temporal dynamics of musical emotions examined through intersubject synchrony of brain activity. *Social Cognitive and Affective Neuroscience*, *10*(12), 1705–1721. https://doi.org/10.1093/scan/nsv060

Turner, J. C. (1991). *Social influence (pp. Xvi, 206)*. New York: Thomson Brooks/Cole Publishing Co.

Uchino, B. N., Smith, T. W., Birmingham, W., & Carlisle, M. (2011). Social neuroscientific pathways linking social support to health. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford Handbook of Social Neuroscience* (pp. 804–813). New York: Oxford University Press.

Umberson, D., & Karas Montez, J. (2010). Social relationships and health: A flashpoint for health policy. *Journal of Health and Social Behavior*, *51*, 54–66. https://doi.org/10.1177/0022146510383501

van Baar, J. M., Halpern, D. J., & FeldmanHall, O. (2021). Intolerance of uncertainty modulates brain-to-brain synchrony during politically polarized perception. *Proceedings of the National Academy of Sciences*, *118*(20), e2022491118. https://doi.org/10.1073/pnas.2022491118

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard. *Journal of Cognitive Neuroscience, 23*(11), 3343–3354. https://doi.org/10.1162/jocn_a_00016

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science, 19*, 1131–1139. doi: 10.1111/j.1467-9280.2008.02214.x.

Van der Meij, L., Almela, M., Hidalgo, V., Villada, C., Ijzerman, H., van Lange, P. A., & Salvador, A. (2012). Testosterone and cortisol release among Spanish soccer fans watching the 2010 World Cup final. *PLOS ONE, 7*(4), e34814. https://doi.org/10.1371/journal.pone.0034814

Van Elk, M., & Snoek, L. (2020). The relationship between individual differences in gray matter volume and religiosity and mystical experiences: A preregistered voxel-based morphometry study. *European Journal of Neuroscience, 51*(3), 850-865. https://doi.org/10.1111/ejn.14563

Vanman, E. J., Kappas, A., & Ito, T. A. (2021). Social neuroscience is more than the study of the human brain: The legacy of John Cacioppo. *Social Neuroscience*, *16*(1), 1–5. https://doi.org/10.1080/17470919.2021.1879459

Veltkamp, M., Aarts, H., & Custers, R. (2008). Perception in the service of goal pursuit: Motivation to attain goals enhances the perceived size of goal-instrumental objects. *Social Cognition*, *26*(6), 720–736. https://doi.org/10.1521/soco.2008.26.6.720

Vora, A., Nguyen, A. D., Spicer, C., & Li, W. (2022). The impact of social isolation on health and behavior in Drosophila melanogaster and beyond. *Brain Science Advances, 8*(3), 183-196. https://doi.org/10.26599/BSA.2022.9050016

Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*(4), 1048–1056. https://doi.org/10.1016/j.jesp.2007.10.009

Walsh, A. (1997). Methodological Individualism and Vertical Integration in the Social Sciences. *Behavior and Philosophy, 25*(2), 121–136. https://www.jstor.org/stable/27759372

Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., ... & Fan, J. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *Elife, 8*, e42265. https://doi.org/10.7554/eLife.42265

Wang, Y., John, D. R., & Griskevicius, V. (2015). The devil wears prada? How luxury consumption influences social behaviors. *ACR North American Advances*, *3*, 142–146. https://doi.org/10.1016/j.ijresmar.2020.04.001

Watanabe, N., & Yamamoto, M. (2015). Neural mechanisms of social dominance. *Frontiers in Neuroscience*, *9*, 154. https://doi.org/10.3389/fnins.2015.00154

Watanabe, T., Sasaki, Y., Shibata, K., & Kawato, M. (2017). Advances in fMRI real-time neurofeedback. *Trends in Cognitive Sciences*, *21*(12), 997-1010. https://doi.org/10.1016/j.tics.2017.09.010

Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress Contagion: Physiological Covariation Between Mothers and Infants. *Psychological Science*, *25*(4), 934–942. https://doi.org/10.1177/0956797613518352

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*(5), 806–820. https://doi.org/10.1037/0022-3514.39.5.806

Weinstein, N. D., & Klein, W. M. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health Psychology*, *14*(2), 132–140. https://doi.org/10.1037/0278-6133.14.2.132

West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology*, *17*(16), 661–672. https://doi.org/10.1016/j.cub.2007.06.004

Wheatley, T., Kang, O., Parkinson, C., & Looser, C. E. (2012). From mind perception to mental connection: Synchrony as a mechanism for social understanding. *Social and Personality Psychology Compass*, *6*(8), 589–606. https://doi.org/10.1111/j.1751-9004.2012.00450.x

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, *16*(1), 56–63. https://doi.org/10.1111/j.0956-7976.2005.00780.x

Wiggins, J. S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behavior: Evidence concerning some untested assumptions underlying diagnostic classification.

Journal of Personality and Social Psychology, *56*(2), 296–305. https://doi.org/10.1037/0022-3514.56.2.296

Wilkinson, R. G., & Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine*, *62*(7), 1768–1784. https://doi.org/10.1016/j.socscimed.2005.08.03s

Wilkinson, R. G., & Pickett, K. E. (2018). *The Inner Level*. Penguin Books.

Wohltjen, S., Toth, B., Boncz, A., & Wheatley, T. (2023). Synchrony to a beat predicts synchrony with other minds. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-29776-6

World Bank (2014). Ending poverty and sharing prosperity. https://www.worldbank.org/content/dam/Worldbank/gmr/gmr2014/GMR_2014_Full_Report. DOI: 10.1596/978-1-4648-0336-9

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLOS ONE*, 8(7), e68910. https://doi.org/10.1371/journal.pone.0068910

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, *29*(26), 8525–8529. https://doi.org/10.1523/JNEUROSCI.2418-09.2009

Yang, C. L., Zhang, B., Charness, G., Li, C., & Lien, J. W. (2018). Endogenous rewards promote cooperation. *Proceedings of the National Academy of Sciences, 115*(40), 9968-9973. https://doi.org/10.1073/pnas.1808241115

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665-670. https://doi.org/10.1038/nmeth.1635

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: Where the idiosyncratic self meets the shared social world. *Nature Reviews Neuroscience*, *22*(3), 181–192. https://doi.org/10.1038/s41583-020-00420-w

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same story, different story: The neural representation of interpretive frameworks. *Psychological Science*, *28*, 307–319. https://doi.org/10.1177/0956797616682029

Yoder, K. J., & Decety, J. (2020). Me first: Neural representations of fairness during three-party interactions. *Neuropsychologia*, *147*, 107576. https://doi.org/10.1016/j.neuropsychologia.2020.107576

Yu, R., Calder, A. J., & Mobbs, D. (2014). Overlapping and distinct representations of advantageous and disadvantageous inequality. *Human Brain Mapping*, *35*(7), 3290–3301. https://doi.org/10.1002/hbm.22402

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychological Science*, *22*(7), 894–900. https://doi.org/10.1177/0956797611411057

Zaki, J., Wager, T. D., Singer, T., Keysers, C., & Gazzola, V. (2016). The Anatomy of Suffering: Understanding the Relationship between Nociceptive and Empathic Pain. *Trends in Cognitive Sciences*, *20*(4), 249–259. https://doi.org/10.1016/j.tics.2016.02.003

Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological Bulletin*, *146*(2), 118–149. https://doi.org/10.1037/bul0000218

Zerubavel, N., Bearman, P. S., Weber, J., & Ochsner, K. N. (2015). Neural mechanisms tracking popularity in real-world social networks. *Proceedings of the National Academy of Sciences*, *112*(49), 15072–15077. https://doi.org/10.1073/pnas.1511477112

Zhang, L., Chen, P., Schafer, M., Zheng, S., Chen, L., Wang, S., Liang, Q., Qi, Q., Zhang, Y., & Huang, R. (2022). A specific brain network for a social map in the human brain. *Scientific Reports*, *12*(1), 1773. https://doi.org/10.1038/s41598-022-05601-4

Zhong, C. B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science*, *21*(3), 311–314. https://doi.org/10.1177/0956797609360754

Zhou, S., Xu, X., He, X., Zhou, F., Zhai, Y., Chen, J., Long, Y., Zheng, L., & Lu, C. (2022). Biasing the neurocognitive processing of videos with the presence of a real cultural other. *Cerebral Cortex*, bhac122. https://doi.org/10.1093/cercor/bhac122

Zink, C. F., Tong, Y., Chen, Q., Bassett, D. S., Stein, J. L., & Meyer-Lindenberg, A. (2008). Know your place: Neural processing of social hierarchy in humans. *Neuron*, *58*(2), 273–283. https://doi.org/10.1038/s41562-017-0045

Zoh, Y., Chang, S. W., & Crockett, M. J. (2022). The prefrontal cortex and (uniquely) human cooperation: A comparative perspective. *Neuropsychopharmacology*, *47*(1), 119–133. https://doi.org/10.1038/s41386-021-01092-5

ENDNOTES

The Handbook of Social Psychology, 6th edition © 2025 by Situational Press is licensed under Creative Commons- Attribution-NonCommercial-NoDerivatives 4.0 International. This work may be copied and distributed only in unmodified form, for noncommercial purposes, and with attribution.