The neuroscience of morality and social decision-making

Keith J. Yoder[a] and **Jean Decety**[a,b,*]

[a]Department of Psychology, University of Chicago, Chicago, IL, USA

[b]Department of Psychiatry and Behavioral Neuroscience, University of Chicago Medicine, Chicago, IL, USA

Abstract

Across cultures humans care deeply about morality and create institutions, such as criminal courts, to enforce social norms. In such contexts, judges and juries engage in complex social decision-making to ascertain a defendant's capacity, blameworthiness, and culpability. Cognitive neuroscience investigations have begun to reveal the distributed neural networks which interact to implement moral judgment and social decision-making, including systems for reward learning, valuation, mental state understanding, and salience processing. These processes are fundamental to morality, and their underlying neural mechanisms are influenced by individual differences in empathy, caring and justice sensitivity. This new knowledge has important implication in legal settings for understanding how triers of fact reason. Moreover, recent work demonstrates how disruptions within the social decision-making network facilitate immoral behavior, as in the case of psychopathy. Incorporating neuroscientific methods with psychology and clinical neuroscience has the potential to improve predictions of recidivism, future dangerousness, and responsivity to particular forms of rehabilitation.

Keywords

Introduction

All members of human societies are characterized by a deep concern over issues of morality, justice, and fairness (Decety & Yoder, 2017). In fact, humans are unique among other living creatures in establishing cultural organizations to enforce particular social norms, including institutions designed to evaluate the acceptability of individuals' behaviors and assign appropriate punishments to those who violate particular norms (Buckholtz & Marois, 2012). Regardless of how morality is conceptualized, many scholars, following Darwin (1871), have made the claim that it is an evolved aspect of human nature that facilitates cooperation in large groups of unrelated individuals (Tomasello & Vaish, 2013). Associating in groups improves the chances of survival compared with solitary existence. Moral norms provide safeguards against infringements on safety or health, and reinforcement of moral behaviors

*Corresponding author: Dr. Jean Decety, Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, IL, USA, decety@uchicago.edu.

minimizes criminal behavior and social conflict. In these ways, morality makes human society possible.

There are at least two reasons to think morality bears the imprint of natural selection. While nonhuman animals obviously don't reason explicitly about right and wrong, good and bad, just and unjust, or vice and virtue, some exhibit behaviors which seem to incorporate elements of human morality. Many species cooperate, help their kin, and care for their offspring (Tremblay, Sharika, & Platt, 2017), and some manifest inequity aversion (Decety & Yoder, 2017. Likewise, while socialization influences moral development and explains why moral rules change with space and historical time, human infants enter the world equipped with cognitions and motivations that incline them to be moral and prosocial (Hamlin, 2015). Such early emerging predispositions toward prosocial behavior, and sociomoral evaluation reflect prewired capacities that were adaptive to our forebears. Members of *homo sapiens* cooperate with and help non-relatives at cost to themselves at a rate that is unmatched in the rest of the animal kingdom. This penchant for cooperation with unrelated individuals explains why our species successfully colonized the entire planet (Marean, 2015). However, this does not imply that morality is itself an adaptation favored by natural selection. Instead, the moral sense observed in humans may be a consequence of several cognitive, executive, and motivational capacities which are the attributes that natural selection directly promoted (Ayala, 2010).

Decades of research across multiple disciplines, including behavioral economics, developmental psychology, and social neuroscience, indicate that moral reasoning arises from complex social decision-making and involves both unconscious and deliberate processes which rely on several partially distinct dimensions, including intention understanding, harm aversion, reward and value coding, executive functioning, and rule learning (Decety & Cowell, 2017; Gray, Young, & Waytz, 2012; Krueger & Hoffman, 2016; Ruff & Fehr, 2014). Human moral decisions are governed by both statistical expectations (based on observed frequencies) about what others will do and normative beliefs about what others should do. These vary across different cultures and historical contexts, forming a continuum from social conventions to moral norms which typically concern harm to others.

In this article, we first discuss recent empirical progress in the neuroscience of social decision-making. Next, we examine the neural mechanisms underlying the components involved in morality. Then, a section is devoted to psychopathy because it constitutes a model for understanding the consequences arising from atypical neural processing and lack of concern for others and moral rules. Finally, we discuss how this social neuroscience perspective has valuable forensic implications, both in terms of understanding how jurors and judges make decisions about culpability and severity, and in predicting defendant's future behavior. Understanding the psychological mechanisms and neurological underpinnings of how we make moral decisions sheds light on the diagnosis and treatment of the serious wrongdoers among us.

Social decision-making

Human lives consist of a constant stream of decisions and choices. Essentially, the study of decision-making attempts to understand the fundamental ability to process multiple alternatives and to choose an optimal course of action (Sanfey, 2007). This involves identifying rewarding stimuli to approach and aversive stimuli to avoid (Pessiglione & Delgado, 2015). However, successful decision-making also requires maintaining representations of short-term and long-term goals in order to maximize decision payoffs. Social decision-making specifically refers to decisions in which individuals consider and integrate the goals, beliefs, and intentions of other individuals into their decision-making calculus. Thus, social decision-making relies on Theory of Mind, the capacity to attribute mental states such as beliefs, intentions and desires to oneself and others. However, social decision-making is also at times constrained by social norms and associated punishments, which affect the expected value of specific response options.

The neural systems supporting social decision-making have been investigated from multiple perspectives, including neuroeconomics, cognitive neuroscience, and translational neuroscience (Figure 1). Converging evidence indicates that the components of social decision-making rely on the coordination of multiple neurocognitive systems which support domain-general processes such as stimulus valuation, perspective-taking, mental state understanding and response selection (Berridge & Kringelbach, 2015; Ruff & Fehr, 2014; Tremblay et a 2017).

In non-social contexts, rewarding and aversive stimuli are associated with activity in largely separate neural networks (Pessiglione & Delgado, 2015). For instance, functional magnetic resonance imaging (fMRI) studies, have shown that the striatum, rostral anterior cingulate cortex (rACC), and ventromedial prefrontal cortex (vmPFC) are reliably recruited in response to rewarding stimuli. In contrast, dorsal ACC (dACC), anterior insula (aINS), and amygdala are more active in response to aversive stimuli including pain. More specifically, experiencing (as opposed to anticipating) a rewarding stimulus (e.g. eating ice cream or winning the lottery) is associated with neural responses in orbitofrontal cortex (OFC), ventromedial prefrontal cortex (vmPFC), and the amygdala (Ruff & Fehr, 2014). Such experience-reward signals are important inputs to distinct neurons in OFC/vmPFC, amygdala, and striatum, which update and maintain the stimulus-value associations and expectations employed in judgment and decision-making (Balleine, Delgado, & Hikosaka, 2007; Berridge & Kringelbach, 2015; Wassum & Izquierdo, 2015). In contrast to rewards, aversive experiences (e.g. pain) are reliably associated with activity in aINS, dACC, and distinct neural populations in the amygdala (Pessiglione & Delgado, 2015). Importantly, aINS and dACC anchor the salience network which functions to increase attention to important changes in the environment and motivate avoidance of dangerous or noxious stimuli (Cunningham & Brosch, 2012; Shackman et al., 2011). Finally, the dorsolateral prefrontal cortex (dlPFC) supports goal maintenance and implementing specific goal-directed behaviors (Ruff & Fehr, 2014).

The brain areas described above are also implicated in social decision-making tasks (Ruff & Fehr, 2014). Though one study with neurological patients suggests that social and non-social

value-based decision-making might be separable at the neural level (Besnard et al., 2016), the extant evidence is inconclusive. Since the most common techniques employed in human functional neuroimaging assess coordinated activity across large populations of neurons, it remains an open question whether social and non-social information is supported by the same neurons, or engage distinct parallel channels. For instance, vmPFC seems implicated in both social cognition and valuation, and this has sparked a debate about its primary functional role (Delgado et al., 2016). From a neuroeconomics perspective, the significance of vmPFC is in supporting the representation of the personal subjective value of stimuli with a single "common currency" (Ruff & Fehr, 2014). Others argue that the vmPFC instead primarily functions to support social cognition by maintaining representations of the self and close others (Delgado et al., 2016). An anatomically intact vmPFC is crucial for typical decision-making in tasks which require individuals to distinguish between different rewards and punishments with varying magnitudes and probabilities, such as the Iowa Gambling Task (Gläscher et al., 2012). Social contexts contain highly dynamic information and representations of other individuals. Thus, the vmPFC involvement in social cognition could arise in response to the processing demands required by such contexts. Moreover, mental state attribution, which is can be a necessary input to social decision-making, relies on a network of interconnected cortical regions including the posterior superior temporal sulcus (pSTS), temporoparietal junction (TPJ), medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC), and precuneus (Decety & Lamm, 2007; Tremblay et al., 2017; Young, Cushman, Hauser, & Saxe, 2007). These regions supporting mental state attribution and dlPFC are coactivated specifically during social decision-making This functional coupling reflects the expected value of decisions and provides a crucial framework for social information such as mental state understanding to influence valuation processes (Decety & Yoder, 2017; Delgado et al., 2016).

Overall, though there are still open questions, cognitive neuroscience and neuroeconomics have provided a solid platform for investigating specific cases of social decision-making that are directly relevant to legal settings. First, research into moral cognition has identified the neural computations and networks important for assessing intentionality and harm, as well as how those systems interact to produce judgments of culpability and blameworthiness (Krueger & Hoffman, 2016). Second, the emerging field of neuroethics has begun to clarify the factors that contribute to antisocial behavior, including potential neuropsychological risk factors for future antisocial behavior (Gaudet, Kerkmans, Anderson, & Kiehl, 2016).

Moral values, empathy and justice motivation

People have moral values. They accept standards that allow their behavior to be judged as either right or wrong, good or evil, praiseworthy or blameworthy. Though particular norms by which moral actions are judged vary to some extent from individual to individual and across geographical locations (although some norms, such as not to kill, not to steal, and to honor one's parents, are widespread and perhaps universal), all cultures demonstrate value judgments concerning human behavior. So important are social norms that in every culture studied to date people are willing to give up some of their own resources to punish individuals that violate social norms, even when they themselves are not directly affected (Henrich et al., 2006). Moreover, third parties evaluate moral wrongness and

blameworthiness by taking into account both an action's outcome and an agent's intention (Buckholtz et al., 2008; Young, Scholz, & Saxe, 2011). In this way, individuals' everyday moral judgments match the foundational legal concepts of *mens rea* and *actus rea*, where harmful intentions and harmful outcomes produce additive effects on condemnation. Thus, third-party judgments necessitate social decision-making capacities to integrate the value of particular outcomes and the understanding of the mental state of relevant agents, the latter relying on social cognitive abilities such as theory of mind.

Work across various academic disciplines has converged on the view that morality arises from the integration of both innate abilities which are shaped by natural selection and deliberative processes that interact with social environments and cultural exposure (Decety & Wheatley, 2015). Throughout most of human history individuals have lived in small groups which facilitated repeated interactions with other individuals (Baumard, André, & Sperber, 2013). Moreover, humans' elaborated social cognitive abilities allow individuals to not only observe the behavior of others and predict future behavior, but also communicate that information to one another. Since individuals have some flexibility in selecting social partners, partner choice exerts strong pressure to at least appear as if one will likely cooperate and support reciprocal interaction (Baumard & Sheskin, 2015). This mutualistic perspective posits that the most efficient psychological mechanism to achieve a reputation as a cooperator is through a genuine moral concern which treats cooperation as intrinsically good (Baumard et al., 2013; Debove, Baumard, & André, 2017). In other words, it is evolutionarily beneficial for humans to genuinely prefer some kinds of moral outcomes because it helps to maintain cooperation.

Another theory of morality rests on the idea that humans are equipped with a moral disposition. This naturalistic view of a moral sense is originally grounded in the *Humean* idea that moral judgments arise from an immediate aversive reaction to perceived or imagined harm to victims (Hume, 1738). These actions are judged as immoral only *after*, and because of, the initial affective reaction. On that perspective, social emotions (e.g., empathy, guilt, shame) play a pivotal role in morality. Such social emotions contribute to fitness in facilitating caring for others and group cohesion.

It is important to note that while empathy is a powerful motivation for prosocial behavior, it should not be equated with morality. The two concepts refer to distinct abilities with partially non-overlapping proximate mechanisms (Decety & Cowell, 2014a). Whereas morality deals with social norms prohibiting and prescribing specific behaviors, empathy is a complex multi-faceted construct that involves perspective-taking, affect sharing, and a motivated concern for other's well-being (Decety & Jackson, 2004). Each of these components are implemented in specific brain systems, and have important implications for moral decision-making and behavior (Decety & Cowell, 2014b). For instance, perspective-taking can be used to adopt the subjective viewpoint of others, and this can facilitate understanding the extent of harm or distress that might be experienced by a victim. Conversely, affective reactions to the plight of another may be foundational for motivating prosocial behaviors as well as moral condemnation (Decety & Cowell, 2017; Patil, Calò, Fornasier, Cushman, & Silani, 2017). But affective sharing may also lead to personal distress, the aversive affect arising in response to others' suffering, and does not necessarily

lead to prosocial behavior (Decety & Lamm, 2009). Furthermore, the degree of these empathic responses are known to be modulated (enhanced or suppressed) by social and contextual factors. For instance, stronger reactions and associated neural responses are elicited when observing the pain of people from the same ethnic group compared with people of another group (Contreras-Huerta, Baker, Reynolds, Batalha, & Cunnington, 2013). Many individuals experience schadenfreude when outgroup members experience misfortune (Cikara, Bruneau, & Saxe, 2011). Furthermore, the responses toward a victim are not always compassionate and are moderated by the perceived responsibility of the victim. Increased neural activation in the mPFC and vmPFC was found for innocent compared with blameworthy victims, as well as decreased functional connectivity between dlPFC and mPFC/ACC, and dlPFC for the latter (Fehse, Silveira, Elvers, & Blautzik, 2015). The recruitment of these regions reflects social evaluation processes related to moral reasoning. This finding is congruent with a study which demonstrated that empathic responses toward a victim also engage brain areas related to social understanding and moral evaluation depending on the perceived intent (intentional or unintentional) of who caused the harm (Decety, Michalska, & Akitsuki, 2008). In this way, the impact of empathic responding on moral judgment functions as a double-edged sword during sentencing, with prosecutors working to elicit concern for victims and cast defendants as out-groups, while defense attorneys work to humanize their clients (Johnson, Hritz, Royer, & Blume, 2016). In both cases, attorneys are leveraging social cognition, including mental state attribution, to shift the social decision-making apparatus of jurors and judges towards selecting a particular response.

Justice motivation encompasses issues of fairness, equity, and equality (Decety & Yoder, 2017). In general, people strive to behave in accordance with justice principles and condemn injustice as both morally wrong and worthy of punishment. Recent evidence suggests that a preference for justice emerges very early in development, coincident with dislike and avoidance of injustice (Cowell & Decety, 2015a; Hamlin, 2014; Sommerville, Schmidt, Yun, & Burns, 2013). However, individuals differ in their tendency to detect and react to justice issues (Baumert & Schmitt, 2016), and these dispositional differences predict altruistic sharing, transgressive behavior, and moral judgments (Decety & Yoder, 2015; Gollwitzer, Schmitt, Schalke, Maes, & Baer, 2005). Importantly, personal involvement alters a situation; sensitivity to self-focused and other-focused justice are distinct and follow different developmental pathways (Bondü, Hannuschke, Elsner, & Gollwitzer, 2016). Whereas other-focused justice sensitivity reflects genuine prosocial concern for the welfare of others, self-focused justice sensitivity includes increased distrust of others' intentions (Baumert & Schmitt, 2016). In fact, high self-focus has been linked to more permissive moral judgments and increased antisocial behavior (Decety & Yoder, 2015; Gollwitzer et al., 2005). Moreover, empathic processes and justice motivation are also linked. For instance, perspective-taking and empathic concern, but not personal distress and affective sharing, are positively related to other-oriented justice motivation (Decety & Yoder, 2015).

Cognitive neuroscience of morality

Converging evidence from functional neuroimaging studies and neurological observations indicates that the same regions implicated in social decision-making play important specific

roles in morality (Figure 1). Specifically, a set of interconnected regions encompassing the vmPFC, OFC, amygdala, TPJ, ACC, aINS, PCC, and dlPFC are reliably engaged across tasks which involve explicit or implicit evaluations of morally-laden stimuli, regardless of whether the outcome of an action affects the participants directly or another individual (Eres, Louis, & Molenberghs, 2017). Neuroscience investigations of third-party punishment judgments, such as those made by jurors (Buckholtz et al., 2008), suggest that intention understanding and harm perception rely on interconnected, but largely distinct neural systems (Krueger & Hoffman, 2016). As discussed above, representations of others' mental states, beliefs, and intentions are supported by pSTS/TPJ, PCC/precuneus, and mPFC (Decety & Lamm, 2007; Young et al., 2007, 2011). These regions are foundational to incorporating intentionality into moral judgments. Conversely, perception of harm relies more on aINS, ACC, and the amygdala, core nodes of the salience network (Droutman, Bechara, & Read, 2015; Hesse et al., 2016), reflecting a neural mechanism for an actions' outcome to influence decision-making. Finally, integrating harm and intent in order to determine an appropriate punishment relies on intact functioning of the central executive network, especially lateral parietal cortex and dlPFC (Buckholtz & Marois, 2012; Decety & Cowell, 2017; Krueger & Hoffman, 2016).

Researchers have also built on social decision-making research to examine how domain-general systems contribute to moral judgments[1] in more commonplace, everyday-like situations (Decety & Yoder, 2015; Shenhav & Greene, 2010; Yoder & Decety, 2014a, 2014b). Neuroimaging work in this vein has begun to address the specific roles of perspective-taking, emotional reactivity, and executive functioning in moral cognition (Avram et al., 2014; Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006; Yoder & Decety, 2014a, 2014b). This approach has also been used to identify differences between typical and psychiatric populations (Yoder, Harenski, Kiehl, & Decety, 2015; Yoder, Lahey, & Decety, 2016), and in developmental neuroscience to probe the emergence of various aspects of moral reasoning (Cowell & Decety, 2015a, 2015b; Decety & Cowell, 2017). More importantly, by asking individuals to evaluate the permissibility of carefully constructed everyday interactions, researchers are able to elucidate how the nervous systems supports the sorts of third-party punishment judgments required of judges and juries (Buckholtz et al., 2008; Krueger & Hoffman, 2016).

In healthy individuals, the pain and distress of others is a powerful cue which motivates prosocial behaviors to alleviate suffering and condemn harmful actions (Decety & Cowell, 2015, 2017). Intentional harmful actions are simple to evaluate as morally wrong because of the congruence between intention and outcome. When the two conflict, such as in the case

---

[1]Much of the literature regarding moral cognition in adults has been dominated by studies using sacrificial dilemmas (Kahane, 2015). Perhaps the most famous dilemma is the Trolley Problem (Thomson, 1985), which asks participants to decide if it is morally permissible to pull a lever in order to divert a trolley onto a secondary track, saving the lives of five strangers on the primary track, but leading to the death of one stranger. Ostensibly, pulling the lever is a utilitarian choice, while condemning such an action is indicative of deontological ethics. Greater neural response in dlPFC and ACC, or vmPFC, PCC, and TPJ, have been interpreted within a dual-process framework, where deontological rules arise from emotional responses, but utilitarian judgments depends on cognitive deliberation. This approach has been used to argue for specific emotional deficits in frontal lesion populations (Koenigs et al., 2007) and psychopathy (Koenigs, Kruepke, Zeier, & Newman, 2012). However, critics argue that the legal implications of this approach have been overstated (Pardo & Patterson, 2016). Sacrificial dilemmas have also been criticized because they rarely present a true utilitarian response option, may not require moral reasoning, and are unlikely to reflect how moral decision-making occurs for most people during their everyday lives (Kahane, Everett, Earp, Farias, & Savulescu, 2015).

of accidental harms, pSTS/TPJ drives dACC to exert top-down control on amygdala response, thus blunting the effect of amygdala signaling on response selection processes carried out by dlPFC (Treadway et al., 2014). At the same time, greater empathic concern for the suffering of a victim results in greater functional connectivity between right aINS and dlPFC, and greater condemnation of the accidental act (Patil et al., 2017).

Less is known about the neural underpinnings of justice motivation, though there have been some preliminary investigations into associations between neural functioning and individual dispositions in justice sensitivity (Yoder & Decety, 2014a, 2014b). One study found that individuals with higher dispositional other-oriented justice motivation showed greater activity in the dlPFC when making explicit moral judgments of harmful and helpful dyadic intera (Yoder & Decety, 2014b). Importantly, justice motivation was also positively related to increased functional connectivity between pSTS/TPJ and dlPFC. As discussed above, neuronal coupling between these regions facilitates the incorporation of mental state understanding into the decision-making process. Thus, it appears that an individual's concern for justice principles influences the extent to which they utilize a perpetrator's intentions to inform their moral judgments. Additionally, neuroeconomics research has long documented the role of the social decision-making network in reactions to monetary inequity (Ruff & Fehr, 2014). Early studies interpreted increased activity within the salience network (e.g. aINS) in response to unfair outcomes as reflecting negative affect (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). However, strong emotional responses are not always present during such tasks, suggesting that rejecting inequity instead reflects a cognitive heuristic (Civai, 2013). This is also consistent with the evolved moral concern posited by mutualism discussed above (Baumard et al., 2013).

Overall, moral judgment arises from coordinated activity between domain-general capacities for perspective-taking, salience processing, executive control, valuation, and social norm compliance (Decety & Yoder, 2017; Krueger & Hoffman, 2016). Understanding how the neural networks supporting these computations interact, and how aspects of empathy and justice motivation influence such interactions, is important for elucidating how the humans who serve on juries make decisions about blameworthiness and culpability (Buckholtz et al., 2008). Research clarifying when emotional and cognitive processes may bias decision-making in particular ways also provides crucial information for appellate courts and legislative bodies who must determine whether and how particular arguments and pieces of evidence should be presented (Johnson et al., 2016). In particular, while empathic processes do play important functions in motivating caring for others and guiding moral judgment in various forms, they are highly sensitive to the social identities of persons, their interpersonal relationships, and social context. However, it has recently been argued that the role of empathy in shaping people's understanding of why harming others is wrong and in producing the relevant motivation is more limited than people think (Decety & Cowell, 2015).

Psychopathy and social decision-making

Certainly, antisocial behavior may arise from specific dysfunctions in the social decision-making and morality pathways. Damage or disruption of the functional connectivity in the

network of neural regions underlying social decision-making have already been used as part of *mens rea* defenses (Farahany, 2015). Neuroanatomical abnormalities such as tumors, injuries, or other forms of brain damage, as well as functional neuroimaging evidence, have been offered to suggest that some criminal defendants lack the capacity to understand right from wrong or to behave in accordance with social norms. In such instances, defense lawyers argue that the defendant's brain is abnormal in a way that reduces the defendant's culpability or that the neuroscience evidence should provide mitigation during sentencing. Thus, moral neuroscience has clear forensic implications when examining individuals who decide to engage in criminal behavior. One complex and controversial condition or mental disorder that disproportionately impacts the criminal justice system is psychopathy. Psychopaths are twenty to twenty-five times more likely than non-psychopaths to be in prison, four to eight times more likely to violently recidivate compared to non-psychopaths, and are resistant to most forms of treatment (Kiehl & Hoffman, 2011).

The absence of moral scruple in the presence of otherwise intact intellect in individuals with psychopathy has fascinated psychiatrists and clinical psychologists for a long time. Psychopathy is a constellation of psychological symptoms, including shallow affect, lack of empathy, guilt and remorse, irresponsibility, and impulsivity. These symptoms typically emerge early in childhood and affect all aspects of an individual's life including relationships with family, friends, school, and work. While the etiology of psychopathy is not well understood, there is a growing body of evidence showing that psychopathy is highly associated with aberrant neuronal activity in specific regions of the brain and atypical anatomical connectivity between specific areas (Kiehl, 2015).

Neuroscience evidence is often used in conjunction with standardized assessments. For instance, the Psychopathy Checklist – Revised (PCL-R) is one of the most widely used instruments in forensic psychiatry, in part, because it has high predictive power for future antisocial behavior (Hare, 2016). Individuals with high PCL-R scores consistently show disruption of neural activity and anatomical integrity in the social decision-making network, including reduced structural connectivity between the amygdala and vmPFC (Motzkin, Newman, Kiehl, & Koenigs, 2011; Wolf et al., 2015), atypical function activity within the amygdala and vmPFC during the evaluation of stimuli depicting moral violations (Harenski, Harenski, Shane, & Kiehl, 2010), and reduced neural responses in vmPFC and periaqueductal gray to the pain and distress cues expressed by others (Decety, Skelly, & Kiehl, 2013). In addition, when viewing morally laden scenarios, there are widespread decreases in functional connectivity seeded in pSTS/TPJ and the amygdala, two computational nodes integral to intention understanding and saliency processing (Yoder, Harenski, et al., 2015).

There is no shortage of evidence that psychopathic individuals lack empathy and concern for the well-being of others (Mahmut, Homewood, & Stevenson, 2008; Marsh et al., 2013; Patrick, 2005). Empathic concern plays an essential role in valuing others' welfare and depends on input from particular physiological, emotional, or motivational processes that seem dysfunctional in psychopathic individuals (Decety & Cowell, 2014b). A functional MRI study conducted with a large sample of inmates with varying levels of psychopathy examined neural responses and functional connectivity under two types of perspective-

taking instructions (Decety, Chen, Harenski, & Kiehl, 2013). Participants were scanned while viewing stimuli depicting bodily injuries. During this time they were instructed to adopt an imagine-self and an imagine-other perspective. Affective perspective taking (i.e., the ability to adopt the subjective point of view of another) is a powerful way to elicit empathy and concern for others (Batson, 2012). Results demonstrate that during the imagine-self perspective, participants with high psychopathy showed a typical neural response within the network involved in empathy for pain, including the anterior insula, anterior midcingulate cortex (aMCC), somatosensory cortex, and amygdala. Conversely, during the imagine-other perspective, psychopaths exhibited an atypical pattern of brain activation and effective connectivity seeded in the anterior insula and amygdala with the vmPFC. The failure to recruit the neural circuits that are typically activated in non-psychopathic individuals during the imagine-other perspective could provide a mechanistic explanation for why psychopathic moral knowledge does not translate directly into moral motivation.

Healthy adults appear to spontaneously attend to the pain and distress of others (Decety & Cowell, 2017). In fact, this motivation is so strong that top-down inhibition of the salience network manifests when such information is task-irrelevant. For instance, adults with low-psychopathy scores exhibit greater activity in dlPFC when reporting whether a harmful or helpful interaction occurred outdoors than whether it was morally wrong (Yoder, Harenski, et al., 2015). Psychopaths did not show this enhanced prefrontal activity, instead showing reductions in neuronal coupling between pSTS/TPJ and aINS, suggesting that these individuals perceive third-person distress as less meaningful.

Behavioral investigations into the influence of psychopathy on moral decision-making have yielded contradictory results, possibly because early studies focused either on judgment (abstract evaluation) or on choices between hypothetical actions; two processes that may rely on different mechanisms. For instance, it was argued that psychopathy was characterized by a failure to distinguish between right and wrong when tested on the classic moral/conventional transgressions task (Blair, Jones, Clark, & Smith, 1995)[2]. However, further investigations with forensic populations found no effect of psychopathy on moral classification accuracy, and even individuals with very high psychopathy scores do understand moral rules and can appropriately identify actions as right and wrong (Aharoni, Sinnott-Armstrong, & Kiehl, 2014). These patterns of results support the view that psychopathic individuals know right from wrong but don't care. One study explored the influence of psychopathic traits on judgment and choice in response to hypothetical scenarios in a non-forensic sample (Tassy, Deruelle, Mancini, Leistedt, & Wicker, 2013). Psychopathy did not predict utilitarian judgments during the evaluation of moral dilemmas, but was positively correlated with utilitarian predictions of future behavior.

---

[2]According to the social domain theory, people across cultures and from an early age distinguish between norms whose violation results in unjust treatment or in harmful consequences to others from those whose violation challenges contextually relative and arbitrary social conventions or norms that structure social interactions (Nucci & Nucci, 1982; Turiel, 1983). Findings from a recent fMRI study suggest a common valence-based decision-making underpinning judgments of both harm/welfare-based and social-conventional social, but also indicate that judgments of different norms are marked by differences in the forms of affect associated with their transgression and relative recruitment of specific computational processes (White et al., 2017).

Despite comparable moral judgments and evaluations, psychopathy is marked by abnormal patterns of neural activity. Indeed, newer evidence suggests that the primary abnormality in psychopathy is disrupted stimulus-value representations which rely on amygdala and striatum functioning (Korponay et al., 2017; Moul, Killcross, & Dadds, 2012). For instance, the antisocial-impulsive dimension of psychopathy, which specifically includes antisocial tendencies, is associated with enlarged striatum and abnormal resting connectivity throughout the cortex (Korponay et al., 2017). Inmates with high PCL-R scores demonstrate abnormally large subjective value signals in the ventral striatum and decreased resting connectivity between vmPFC and ventral striatum, and these effects correlate with previous criminal convictions (Hosking et al., 2017).

Importantly, psychopathic personality traits exist on a spectrum that extends into the general population. In the context of moral cognition, higher levels of psychopathic personality traits are also associated with more permissive moral judgments and decreased amygdala-PFC connectivity when viewing violence (Decety & Yoder, 2015; Yoder, Porges, & Decety, 2015).

Behavioral economics paradigms have also linked psychopathy to reduced prospective, but not retrospective, regret (Baskin-Sommers, Stuppy-Sullivan, & Buckholtz, 2016). In that study, individuals with high psychopathy dispositions expressed similar levels of disappointment when losing money, but their choices were less driven by estimates of potential loses. Notably, retrospective regret modulated the relationship between psychopathy and prior arrests.

Overall, the current evidence from the neuroscience literature on moral reasoning and empathy in psychopathy strongly suggests that informed triers of fact are unlikely to find high psychopathic traits sufficient for a *mens rea* defense. In fact, there is currently no evidence that it would be possible, even in principle, for neuroscience to definitely determine a defendant's mental state at a previous moment in time (Morse, 2003). However, neuroscience evidence can be useful during sentencing to bolster predictions of future behavior. The PCL-R itself is highly predictive of recidivism, but structural and functional neuroimaging can further improve predictions of future behavior (Gaudet et al., 2016). For instance, ACC response in male offenders during an inhibitory control task, prior to release, predicted rearrest over the next three years even better than PCL-R scores (Aharoni et al., 2013). Given the difficult task of assessing future dangerousness, providing judges and juries with reliable information is perhaps the most immediate benefit offered by neuroprediction based on the neuroscience of social decision-making.

Conclusion and forensic applications

Social decision-making capacities in humans have allowed them to achieve unprecedented evolutionary success. Decades of research demonstrate that neurocognitive systems for stimulus valuation, mental state attribution, saliency processing, and goal-related response selection provide the necessary mechanisms for moral reasoning. Disruptions in any of these systems can have devastating consequences for individual and collective welfare, which are often dealt with by the legal system. Moreover, atypical changes in the social aspects of

decision-making are pervasive in many neurological and psychiatric disorders (Ruff & Fehr, 2014). A better understanding of the psychological and neural mechanisms of social decision-making and moral behavior is thus an important goal across social and biological sciences with implications for the law and the criminal justice system. The law regards antisocial acts as arising from the same forces which produce all acts of those whose reason is sufficiently intact to ascribe free will, namely, a conscious decision to violate social norms for which, once apprehended, they must be held responsible (Kiehl & Hoffman, 2014).

Current neuroscience work demonstrates that social decision-making and moral reasoning rely on multiple partially overlapping neural networks which support domain-general processes, such as executive control, saliency processing, perspective-taking, reasoning, and valuation. Neuroscience investigations have contributed to a growing understanding of the role of these process in moral cognition and judgments of blame and culpability, exactly the sorts of judgments required of judges and juries. Dysfunction of these networks can lead to dysfunctional social behavior and a propensity to immoral behavior as in the case of psychopathy. Significant progress has been made in clarifying which aspects of social decision-making network functioning are most predictive of future recidivism. Psychopathy, in particular, constitutes a complex type of moral disorder and a challenge to the criminal justice system. Indeed, despite atypical neural processing in specific brain circuits, these individuals are considered sufficiently rational and presumed to have free will to allow moral choice (Kiehl & Hoffman, 2011). Thus psychopaths cannot be excused for their illegal and immoral actions. While future research could identify biomarkers of sufficiently abnormal moral reasoning or reduced capacity to support a *mens rea* defense, there is currently no neuroscience evidence that would be diagnostically exculpatory in the case of psychopathy. It seems more likely that the neuroscience of decision-making could be applied to identifying individuals for targeted interventions that might prove to be even more effective at reducing future antisocial behavior than incarceration. Finally, while most of the evidence discussed in our paper supports the notion that social decision-making and moral reasoning are implemented by domain-general reward, valuation, motivation and reasoning mechanisms, it is not totally clear whether social and non-social valuation are implemented in similar or distinct neuronal populations, or how areas that are specialized for either social and non-social cognitive functions interact across contexts. Knowing if there is an overlap in neural representations of motivational relevance for social and non-social decision-making is important for both conceptual clarity and for improving interventions aims at rehabilitation. In this way, future investigations into the neural networks underpinning social decision-making can help to characterize specific constellations of biomarkers indicating responsiveness to treatment or reduced capacity, which will increase the effectiveness of the legal judgments and lead to better-informed sentencing decisions.

Acknowledgment:

References

Aharoni E, Sinnott-Armstrong W, & Kiehl KA (2014). What's wrong ? Moral understanding in psychopathic offenders. Journal of Research in Personality, 53, 175–181. 10.1016/j.jrp.2014.10.002 [PubMed: 26379315]

Aharoni E, Vincent GM, Harenski CL, Calhoun VD, Sinnott-Armstrong W, Gazzaniga MS, & Kiehl KA (2013). Neuroprediction of future rearrest. Proceedings of the National Academy of Sciences, 110(15), 6223–6228. 10.1073/pnas.1219302110

Avram M, Hennig-Fast K, Bao Y, Poppel E, Reiser M, Blautzik J,… Gutyrchik E (2014). Neural correlates of moral judgments in first- and third-person perspectives: implications for neuroethics and beyond. BMC Neuroscience, 15, 39 10.1186/1471-2202-15-39 [PubMed: 24742205]

Ayala FJ (2010). The difference of being human: Morality. Proceedings of the National Academy of Sciences, 107(Supplement_2), 9015–9022. 10.1073/pnas.0914616107

Balleine BW, Delgado MR, & Hikosaka O (2007). The role of the dorsal striatum in reward and decision-making. The Journal of Neuroscience, 27(31), 8161–5. 10.1523/JNEUROSCI. 1554-07.2007 [PubMed: 17670959]

Baskin-Sommers A, Stuppy-Sullivan AM, & Buckholtz JW (2016). Psychopathic individuals exhibit but do not avoid regret during counterfactual decision making. Proceedings of the National Academy of Sciences, 113(50), 14438–14443. 10.1073/pnas.1609985113

Batson CD (2012). The empathy-altruism hypothesis: Issues and implications In Decety J(Ed.), Empathy - From Bench to Bedside (pp. 41–54). Cambridge, MA: MIT press.

Baumard N, André J-B, & Sperber D (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. Behavioral and Brain Sciences, 36(1), 59–78 10.1017/S0140525X11002202 [PubMed: 23445574]

Baumard N, & Sheskin M (2015). Partner choice and the evolution of a contractualist morality In Decety J & Wheatley T (Eds.), The Moral Brain (pp. 35–48). Cambridge, MA: MIT Press.

Baumert A, & Schmitt M (2016). Justice Sensitivity In Sabbagh C & Schmitt M (Eds.), Handbook of Social Justice Theory and Research (pp. 161–180).New York, NY: Springer10.1007/978-1-4939-3216-0

Berridge KC, & Kringelbach ML (2015). Pleasure systems in the brain. Neuron, 86(3), 646–664. 10.1016/j.neuron.2015.02.018 [PubMed: 25950633]

Besnard J, Le Gall D, Chauviré V, Aubin G, Etcharry-Bouyx F, & Allain P (2016).Discrepancy between social and nonsocial decision-making under uncertainty following prefrontal lobe damage: the impact of an interactionist approach. Social Neuroscience, 12(4), 430–447. 10.1080/17470919.2016.1182066 [PubMed: 27109748]

Blair RJR, Jones L, Clark F, & Smith M (1995). Is the Psychopath "morally insane"? Personality and Individual Differences, 19(5), 741–752. 10.1016/0191-8869(95)00087-M

Bondü R, Hannuschke M, Elsner B, & Gollwitzer M (2016). Inter-individual stabilization of justice sensitivity in childhood and adolescence. Journal of Research in Personality, 64, 11–20. 10.1016/j.jrp.2016.06.021

Borg JS, Hynes C, Van Horn J, Grafton S, & Sinnott-Armstrong W (2006). Consequences, action, and intention as factors in moral judgments: An FMRI investigation. Journal of Cognitive Neuroscience, 18(5), 803–17. 10.1162/jocn.2006.18.5.803 [PubMed: 16768379]

Buckholtz JW, Asplund CL, Dux PE, Zald DH, Gore JC, Jones OD, & Marois R (2008). The neural correlates of third-party punishment. Neuron, 60(5), 930–40.10.1016/j.neuron.2008.10.016 [PubMed: 19081385]

Buckholtz JW, & Marois R (2012). The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. Nature Neuroscience, 15(5), 655–61. 10.1038/nn.3087 [PubMed: 22534578]

Cikara M, Bruneau EG, & Saxe RR (2011). Us and them: Intergroup failures of empathy. Current Directions in Psychological Science, 20(3), 149–153. 10.1177/0963721411408713

Civai C (2013). Rejecting unfairness: Emotion-driven reaction or cognitive heuristic? Frontiers in Human Neuroscience, 7, 126 10.3389/fnhum.2013.00126 [PubMed: 23576973]
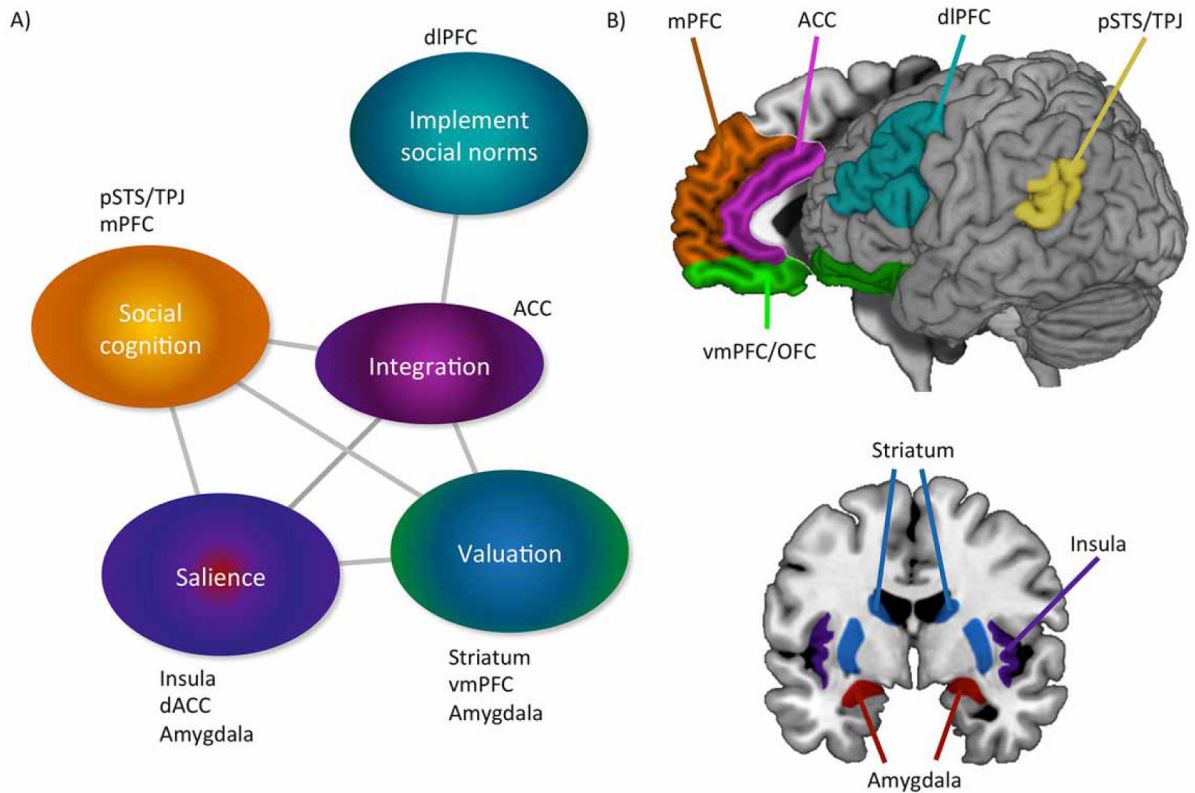
Contreras-Huerta LS, Baker KS, Reynolds KJ, Batalha L, & Cunnington R (2013). Racial bias in neural empathic responses to pain. PLoS ONE, 8(12), e84001 10.1371/journal.pone.0084001 [PubMed: 24376780]

Cowell JM, & Decety J (2015a). Precursors to morality in development as a complex interplay between neural, socioenvironmental, and behavioral facets. Proceedings of the National Academy of Sciences, 112(41), 12657–12662. 10.1073/pnas.1508832112

Cowell JM, & Decety J (2015b). The neuroscience of implicit moral evaluation and its relation to generosity in early childhood. Current Biology, 25(1), 93–97. 10.1016/j.cub.2014.11.002 [PubMed: 25532892]

Cunningham WA, & Brosch T (2012). Motivational salience: amygdala tuning from traits, needs, values, and goals. Current Directions in Psychological Science, 21(1), 54–59. 10.1177/0963721411430832

Darwin C (1871). The Descent of Man and Selection in Relation to Sex (Vol. 1). London, UK Murray.

Debove S, Baumard N, & Andre JB (2017). On the evolutionary origins of equity. PLoS ONE, 12, e0173636 10.1371/journal.pone.0173636 [PubMed: 28323830]

Decety J, Chen C, Harenski C, & Kiehl KA (2013). An fMRI study of affective perspective taking in individuals with psychopathy: Imagining another in pain does not evoke empathy. Frontiers in Human Neuroscience, 7, 489 10.3389/fnhum.2013.00489 [PubMed: 24093010]

Decety J, & Cowell JM (2014a). Friends or foes : Is empathy necessary for moral behaviour? Perspectives on Psychological Science, 9(5), 525–537. 10.1177/1745691614545130 [PubMed: 25429304]

Decety J, & Cowell JM (2014b). The complex relation between morality and empathy. Trends in Cognitive Sciences, 18(7), 337–339. 10.1016/j.tics.2014.04.008 [PubMed: 24972506]

Decety J, & Cowell JM (2015). Empathy, justice, and moral behavior. AJOB Neuroscience, 6(3), 3–14. 10.1080/21507740.2015.1047055 [PubMed: 26877887]

Decety J, & Cowell JM (2017). Interpersonal harm aversion as a necessary foundation for morality: A developmental neuroscience perspective. Development and Psychopathology, 1–12. 10.1017/S0954579417000530

Decety J, & Jackson PL (2004). The functional architecture of human empathy. Behavioral and Cognitive Neuroscience Reviews, 3(2), 71–100. 10.1177/1534582304267187 [PubMed: 15537986]

Decety J, & Lamm C (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. The Neuroscientist, 13(6), 580–93. 10.1177/1073858407304654 [PubMed: 17911216]

Decety J, & Lamm C (2009). Empathy versus personal ditress: Recent evidence from social neuroscience In Decety J & Ickes W (Eds.), The Social Neuroscience of Empathy2 (pp. 199–214). Cambridge, MA: MIT Press.

Decety J, Michalska KJ, & Akitsuki Y (2008). Who caused the pain? An fMRI investigation of empathy and intentionality in children. Neuropsychologia, 46(11), 2607–14. 10.1016/j.neuropsychologia.2008.05.026 [PubMed: 18573266]

Decety J, Skelly LR, & Kiehl KA (2013). Brain response to empathy-eliciting scenarios involving pain in incarcerated individuals with psychopathy. JAMA Psychiatry, 70(6), 638–45. 10.1001/jamapsychiatry.2013.27. [PubMed: 23615636]

Decety J, & Wheatley T (2015). The Moral Brain: A Multidisciplinary Perspective. Cambridge, MA: MIT Press

Decety J, & Yoder KJ (2015). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. Social Neuroscience, 11(1), 1–14. 10.1080/17470919.2015.1029593 [PubMed: 25768232]

Decety J, & Yoder KJ (2017). The emerging social neuroscience of justice motivation. Trends in Cognitive Sciences, 21(1), 6–14. 10.1016/j.tics.2016.10.008 [PubMed: 27865787]

Delgado MR, Beer JS, Fellows LK, Huettel SA, Platt ML, Quirk GJ, & Schiller D (2016). Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex. Nature Neuroscience, 19(12), 1545–1552. 10.1038/nn.4438 [PubMed: 27898086]

Droutman V, Bechara A, & Read SJ (2015). Roles of the different sub-regions of the insular cortex in various phases of the decision-making process. Frontiers in Behavioral Neuroscience, 9, 1–14. 10.3389/fnbeh.2015.00309 [PubMed: 25653603]

Eres R, Louis WR, & Molenberghs P (2017). Common and distinct neural networks involved in fMRI studies investigating morality: An ALE meta-analysis. Social Neuroscience, 0(0), 17470919.2017.1357657. 10.1080/17470919.2017.1357657

Farahany NA (2015). Neuroscience and behavioral genetics in US criminal law: an empirical analysis. Journal of Law and the Biosciences, 2(3), 485–509. 10.1093/jlb/lsv059 [PubMed: 27774210]

Fehse K, Silveira S, Elvers K, & Blautzik J (2015). Compassion, guilt and innocence: An fMRI study of responses to victims who are responsible for their fate. Social Neuroscience, 10(3),243–252. 10.1080/17470919.2014.980587 [PubMed: 25398075]

Gaudet LM, Kerkmans JP, Anderson NE, & Kiehl KA (2016). Can neuroscience help predict future antisocial behavior ? Fordham Law Review, 85(2), 503–531.

Glascher J, Adolphs R, Damasio H, Bechara A, Rudrauf D, Calamia M,… Tranel D (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proceedings of the National Academy of Sciences, 109(36), 14681–6. 10.1073/pnas.1206608109

Gollwitzer M, Schmitt MJ, Schalke R, Maes J, & Baer A (2005). Asymmetrical effects of justice sensitivity perspectives on prosocial and antisocial behavior. Social Justice Research, 18(2), 183–201. 10.1007/s11211-005-7368-1

Gray K, Young L, & Waytz A (2012). Mind perception is the essence of morality. Psychological Inquiry, 23(2), 101–124. 10.1080/1047840X.2012.651387 [PubMed: 22754268]

Hamlin JK (2014). The origins of human morality: Complex socio-moral evaluations by preverbal infants In Decety J & Christen Y (Eds.), New Frontiers in Social Neuroscience (pp. 165–188). New York, NY: Springer.

Hamlin JK (2015). The infantile origins of our moral brains In Decety J & Wheatley T (Eds.), The Moral Brain: A Multidisciplinary Perspective (pp. 105–122). Cambridge, MA: MIT Press

Hare RD (2016). Psychopathy, the PCL-R, and criminal justice: Some new findings and current issues. Canadian Psychology, 57(1), 21–34. 10.1037/cap0000041

Harenski CL, Harenski KA, Shane MS, & Kiehl KA (2010). Aberrant neural processing of moral violations in criminal psychopaths. Journal of Abnormal Psychology, 119(4), 863–74. 10.1037/a0020979 [PubMed: 21090881]

Henrich J, McElreath R, Barr A, Ensminger J, Barrett C, Bolyanatz A, … Ziker J (2006). Costly punishment across human societies. Science, 312(5781), 1767–1770. 10.1126/science.1127333 [PubMed: 16794075]

Hesse E, Mikulan E, Decety J, Sigman M, Garcia M del C, Silva W, … Ibáñez A (2016). Early detection of intentional harm in the human amygdala. Brain, 139, 54–61. 10.1093/brain/awv336 [PubMed: 26608745]

Hosking JG, Kastman EK, Dorfman HM, Samanez-Larkin GR, Baskin-Sommers A, Kiehl KA, … Buckholtz JW (2017). Disrupted prefrontal regulation of striatal subjective value signals in psychopathy. Neuron, 95(1), 221–231. 10.1016/j.neuron.2017.06.030 [PubMed: 28683266]

Hume D (1738). A Treatise on Human Nature. (Norton DF & Norton MJ, Eds.). Oxford, UK: Clarendon Press, 2000.

Johnson SL, Hritz AC, Royer CE, & Blume JH (2016). When empathy bites back: Cautionary tales from neuroscience for capital sentencing. Fordham Law Review, 85, 573–598. 10.3868/s050-004-015-0003-8

Kahane G (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. Social Neuroscience, 10(5), 551–560. 10.1080/17470919.2015.1023400 [PubMed: 25791902]

Kahane G, Everett JAC, Earp BD, Farias M, & Savulescu J (2015). "Utilitarian" judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. Cognition, 134, 193–209. 10.1016/j.cognition.2014.10.005 [PubMed: 25460392]

Kiehl KA, & Hoffman MB (2011). The criminal psychopath: History, neuroscience, treatment, and economics. Jurimetrics, 51, 355–397. 10.1108/17506200710779521 [PubMed: 24944437]

Koenigs M, Kruepke M, Zeier J, & Newman JP (2012). Utilitarian moral judgment in psychopathy. Social Cognitive and Affective Neuroscience, 7(6), 708–14. 10.1093/scan/nsr048 [PubMed: 21768207]

Koenigs M, Young L, Adolphs R, Tranel D, Cushman F, Hauser M, & Damasio A (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. Nature, 446(7138), 908–11. 10.1038/nature05631 [PubMed: 17377536]

Korponay C, Pujara M, Deming P, Philippi C, Decety J, Kosson DS, … Koenigs M (2017). Impulsive-Antisocial dimension of psychopathy linked to enlargement and abnormal functional connectivity of the striatum. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2(2), 149–157. 10.1016/j.bpsc.2016.07.004 [PubMed: 28367514]

Krueger F, & Hoffman M (2016). The emerging neuroscience of third-party punishment. Trends in Neurosciences, 39(8), 499–501. 10.1016/j.tins.2016.06.004 [PubMed: 27369844]

Mahmut MK, Homewood J, & Stevenson RJ (2008). The characteristics of non-criminals with high psychopathy traits: Are they similar to criminal psychopaths? Journal of Research in Personality, 42(3), 679–692. http://doi.Org/10.1016/j.jrp.2007.09.002

Marean CW (2015). The most invasive species of all. Scientific American, 313(2), 32–39 10.1038/scientificamerican0815-32

Marsh AA, Finger EC, Fowler KA, Adalio CJ, Jurkowitz ITN, Schechter JC, … Blair RJR (2013). Empathic responsiveness in amygdala and anterior cingulate cortex in youths with psychopathic traits. Journal of Child Psychology and Psychiatry, 54(8), 900–910 10.1111/jcpp.12063 [PubMed: 23488588]

Morse SJ (2003). Inevitable mens rea. Harvard Journal of Law & Public Policy, 27, 51–64.

Motzkin JC, Newman JP, Kiehl KA, & Koenigs M (2011). Reduced prefrontal connectivity in psychopathy. The Journal of Neuroscience, 31(48), 17348–17357. 10.1523/JNEUROSCI.4215-11.2011 [PubMed: 22131397]

Moul C, Killcross S, & Dadds MR (2012). A model of differential amygdala activation in psychopathy. Psychological Review, 119(4), 789–806. 10.1037/a0029342 [PubMed: 22800411]

Nucci LP, & Nucci MS (1982). Children's responses to moral and social conventional transgressions in free-play settings. Child Development, 53(5), 1337–1342.

Pardo MS, & Patterson D (2016). The promise of neuroscience for law: "Overclaiming" in jurisprudence, morality, and economics In Patterson D & Pardo MS (Eds.), Philosophical Foundations of Law and Neuroscience (pp. 231–248). Oxford: Oxford University Press.

Patil I, Calò M, Fornasier F, Cushman F, & Silani G (2017). The neural basis of empathic blame. Scientific Reports, 7(5200), 1–14. 10.1038/s41598-017-05299-9 [PubMed: 28127051]

Patrick CJ (2005). Handbook of Psychopathy. New York, NY: Guilford Press.

Pessiglione M, & Delgado MR (2015). The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. Current Opinion in Behavioral Sciences, 5, 78–84. http://doi.Org/10.1016/j.cobeha.2015.08.006

Ruff CC, & Fehr E (2014). The neurobiology of rewards and values in social decision making. Nature Reviews Neuroscience, 15(8), 549–562. 10.1038/nrn3776 [PubMed: 24986556]

Sanfey AG (2007). Social decision-making: Insights from game theory and neuroscience. Science, 318(5850), 598–602. 10.1126/science.1142996 [PubMed: 17962552]

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, & Cohen JD (2003). The neural basis of economic decision-making in the Ultimatum Game. Science, 300(5626), 1755–1758. 10.1126/science.1082976 [PubMed: 12805551]

Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, & Davidson RJ (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. Nature Reviews Neuroscience, 12(3), 154–67. 10.1038/nrn2994 [PubMed: 21331082]

Shenhav A, & Greene JD (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. Neuron, 67(4), 667–677. 10.1016/j.neuron.2010.07.020 [PubMed: 20797542]

Sommerville JA, Schmidt MFH, Yun J, & Burns M (2013). The development of fairness expectations and prosocial behavior in the second year of life. Infancy, 18(1), 40–66. 10.1111/j.1532-7078.2012.00129.x

Tassy S, Deruelle C, Mancini J, Leistedt S, & Wicker B (2013). High levels of psychopathic traits alters moral choice but not moral judgment. Frontiers in Human Neuroscience, 7(June), 229 10.3389/fnhum.2013.00229 [PubMed: 23761743]

Thomson JJ (1985). The trolley problem. The Yale Law Journal, 94(6), 1395–1415. 10.1119/1.1976413

Treadway MT, Buckholtz JW, Jan K, Asplund CL, Ginther MR, … Marois R (2014). Corticolimbic gating of emotion-driven punishment. Nature Neuroscience, 17(9), 1270–1275. 10.1038/nn.3781 [PubMed: 25086609]

Tremblay S, Sharika KM, & Platt ML (2017). Social decision-making and the brain: A comparative perspective. Trends in Cognitive Sciences, 21(4), 265–276. 10.1016/j.tics.2017.01.007 [PubMed: 28214131]

Turiel E (1983). The Development of Social Knowledge: Morality and Social Convention. Cambridge, UK: Cambridge University Press.

Wassum KM, & Izquierdo A (2015). The basolateral amygdala in reward learning and addiction. Neuroscience & Biobehavioral Reviews, 57, 10.1016/j.neubiorev.2015.08.017

White SF, Zhao H, Leong KK, Smetana JG, Nucci LP, & Blair RJR (2017). Neural correlates of conventional and harm/welfare-based moral decision-making. Cognitive, Affective, & Behavioral Neuroscience, 1–15. 10.3758/s13415-017-0536-6

Wolf RC, Pujara MS, Motzkin JC, Newman JP, Kiehl KA, Decety J, … Koenigs M (2015). Interpersonal traits of psychopathy linked to reduced integrity of the uncinate fasciculus. Human Brain Mapping, 36(10), 4202–4209. 10.1002/hbm.22911 [PubMed: 26219745]

Yoder KJ, & Decety J (2014a). Spatiotemporal neural dynamics of moral judgment: A high-density ERP study. Neuropsychologia, 60, 39–45. 10.1016/j.neuropsychologia.2014.05.022 [PubMed: 24905282]

Yoder KJ, & Decety J (2014b). The good, the bad, and the just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. Journal of Neuroscience, 34(12), 4161–4166. 10.1523/JNEUROSCI.4648-13.2014 [PubMed: 24647937]

Yoder KJ, Harenski C, Kiehl KA, & Decety J (2015). Neural networks underlying implicit and explicit moral evaluations in psychopathy. Translational Psychiatry, 5(8), e625 10.1038/tp.2015.117 [PubMed: 26305476]

Yoder KJ, Lahey BB, & Decety J (2016). Callous traits in children with and without conduct problems predict reduced connectivity when viewing harm to others. Scientific Reports, 6(February), 20216 10.1038/srep20216 [PubMed: 26832606]

Yoder KJ, Porges EC, & Decety J (2015). Amygdala subnuclei connectivity in response to violence reveals unique influences of individual differences in psychopathic traits in a nonforensic sample. Human Brain Mapping, 36, 1417–1428. 10.1002/hbm.22712 [PubMed: 25557777]

Young L, Cushman F, Hauser M, & Saxe R (2007). The neural basis of the interaction between theory of mind and moral judgment. Proceedings of the National Academy of Sciences of the United States of America, 104(20), 8235–40. 10.1073/pnas.0701408104 [PubMed: 17485679]

Young L, Scholz J, & Saxe R (2011). Neural evidence for "intuitive prosecution": the use of mental state information for negative moral verdicts. Social Neuroscience, 6(3), 302–15. 10.1080/17470919.2010.529712 [PubMed: 21271462]

Figure 1:

Cognitive architecture and brain regions underlying social decision-making and morality. Schematic diagram (A) and color-coded cortical and subcortical areas (B) with their respective roles in decision-making. The salience network is anchored by reciprocal connections between the amygdala, anterior insula, and dorsal anterior cingulate cortex (dACC). It coordinates widespread shifts in neural recruitment in response to motivationally relevant cues. The ventral striatum, amygdala, ventromedial prefrontal cortex (vmPFC), and orbitofrontal cortex (OFC) update and maintain stimulus-value associations, which are essential to reward learning. The posterior superior temporal sulcus (pSTS), temporoparietal junction (TPJ), and mPFC are core nodes underlying social cognitive functions, especially mental state understanding. The ACC is an integrative hub which receives inputs from these diverse regions and is critically involved in computing the anticipated reward value of alternative actions, particularly in situations where action-outcome contingencies vary. The dorsolateral prefrontal cortex (dlPFC) dynamically contributes to cognitive control and instigating goal-direct behaviors. In the context of social decision-making, dlPFC is critical for implementing social norms.