Neural foundations to moral reasoning and antisocial behavior

Adrian Raine and Yaling Yang
Department of Psychology and Neuroscience Program, University of Southern California, USA

A common feature of the antisocial, rule-breaking behavior that is central to criminal, violent and psychopathic individuals is the failure to follow moral guidelines. This review summarizes key findings from brain imaging research on both antisocial behavior and moral reasoning, and integrates these findings into a neural moral model of antisocial behavior. Key areas found to be functionally or structurally impaired in antisocial populations include dorsal and ventral regions of the prefrontal cortex (PFC), amygdala, hippocampus, angular gyrus, anterior cingulate and temporal cortex. Regions most commonly activated in moral judgment tasks consist of the polar/medial and ventral PFC, amygdala, angular gyrus and posterior cingulate. It is hypothesized that the rule-breaking behavior common to antisocial, violent and psychopathic individuals is in part due to impairments in some of the structures (dorsal and ventral PFC, amygdala and angular gyrus) subserving moral cognition and emotion. Impairments to the emotional component that comprises the feeling of what is moral is viewed as the primary deficit in antisocials, although some disruption to the cognitive and cognitive-emotional components of morality (particularly self-referential thinking and emotion regulation) cannot be ruled out. While this neurobiological predisposition is likely only one of several biosocial processes involved in the etiology of antisocial behavior, it raises significant moral issues for the legal system and neuroethics.

Keywords: antisocial; psychopathy; moral; prefrontal; temporal

INTRODUCTION
The burgeoning field of social neuroscience is beginning to provide important insights into the neural mechanisms that underlie the cognitive and affective processes that guide social behavior in everyday life. One particularly important sub-field within this area that has significant societal implications concerns the neural basis to antisocial behavior. The perspective that will be developed here is that there are some similarities between the neural system underlying moral decision-making in normal individuals, and brain mechanisms thought to be impaired in delinquent, criminal, violent and psychopathic populations. We suggest that this is not a chance association but instead represents a neural insight into the etiology of antisocial behavior.

A basic tenet underlying this thesis is that despite some differences in the constructs of childhood conduct disorder, antisocial personality disorder (APD), violence and psychopathy, a critically important common denominator to all is the failure to conform to the commonly accepted mores of society. Clearly, not all ‘immoral’ behavior is illegal, and not all features of a psychopathic personality pertain to immorality; there is more to antisocial behavior than a breakdown in the neural networks subserving moral thinking and feeling. Nevertheless, the overlap between morality and antisocial disorders is substantial, and it is argued here that this is partly accounted for by disruption to neural systems common to both.

A general overview of studies of the neural structures impaired in those with persistent and significant antisocial behavior will first be outlined. The developing knowledge-base on neural mechanisms underpinning moral judgment will then be presented, with an emphasis on the best-replicated neural correlates of moral decision-making. Differences and similarities between moral and antisocial neural correlates will be delineated, and a hypothetical description provided of how disruption to both affective and cognitive components of morality may predispose to the rule-breaking that is central to antisocial disorders.

IMAGING FINDINGS ON ANTISOCIAL, VIOLENT AND PSYCHOPATHIC GROUPS
Despite the increasing evidence for neurological impairment in antisocial individuals, very few structural and functional brain imaging studies have been conducted specifically on the recognized medical disorder for antisocial behavior, i.e. APD. For this reason, the following review will include brain imaging studies on antisocial, violent and psychopathic populations as well as those with APD. While these groups make up different populations, the key thesis of this study is that a disruption to the neural systems underlying moral thinking and feeling gives rise to an ‘antisocial tendency’ (rule-breaking behavior), a key common facet to all antisocial conditions. We group the structural and functional imaging findings based...
on anatomical structures as follows: frontal lobe [orbitofrontal cortex (OFC)/dorsolateral prefrontal cortex (DLPFC)], temporal lobe (superior temporal gyrus/amygdala/hippocampus) and other brain areas (parietal lobe/angular gyrus/cingulate cortex).

**Frontal lobe**

**Global prefrontal abnormalities.** Prefrontal impairments are, perhaps, the best-replicated finding in the imaging literature on antisocial behavior. Reduced prefrontal glucose metabolism has been observed in murderers compared with normal controls (Raine et al., 1994), while acts of aggressive impulsive behavior have been associated with reduced metabolism in the orbitofrontal, anterior medial frontal, and left anterior frontal regions (Goyer et al., 1994). Using single photon emission computed tomography (SPECT), several studies have found significant correlations between reduced frontal blood flow and increased antisocial, aggressive behaviors (Oder et al., 1992; Kuruoglu et al., 1996; Gerr a et al., 1998; Soderstrom et al., 2000, 2002).

These functional impairments are parallel by some evidence for structural prefrontal impairments. Three anatomical MRI (aMRI) studies have found significantly reduced prefrontal gray matter in antisocial and psychopathic individuals (Raine et al., 2000; Yang et al., 2005) and aggressive patients with temporal lobe epilepsy (Woermann et al., 2000), although there is also one null finding (Dolan et al., 2002). Similar findings were reported by Kruesi et al., (2004) where a large (16%) reduction in prefrontal gray volume was found in children diagnosed with conduct disorder. However, given the small sample size ($N=10$), this finding was not statistically significant. Although these findings suggest widespread prefrontal deficits in APD, new evidence has begun to accumulate indicating that the abnormality in APD may be localized to the orbitofrontal and dorsolateral prefrontal regions.

**Orbitofrontal cortex (OFC).** Neurological research has indicated that individuals with orbitofrontal lesions are typically disinhibited, impulsive and unconcerned with the consequences of their behavior (Rolls et al., 1994; Brower and Price, 2001). However, only two aMRI studies to date show reduced gray matter volumes in the OFC in antisocial individuals (Laakso et al., 2002; Yang et al., 2006a). Regarding function, two positron emission tomography (PET) studies have found reduced glucose metabolism in the OFC and medial PFC in impulsive patients and aggressive children (Juhasz et al., 2001; Siever et al., 1999). Similar findings were reported in another two PET studies showing significant negative correlations between aggression and metabolism rate in the OFC and medial frontal cortex (Goyer et al., 1994; Pietrini et al., 2000). In addition, several fMRI studies have shown abnormal OFC activation in impulsive individuals during response inhibition (Horn et al., 2003) and in APD patients during both inhibitory control (Vollm et al., 2004) and fear-conditioning tasks (Birbaumer et al., 2005).

**Dorsolateral Prefrontal Cortex (DLPFC).** Patient studies have shown that damage to the DLPFC classically leads to problems in planning, attention shifting, decision-making and perseverative responding (Manes et al., 2002; Gomez-Beldarrain et al., 2004). However, only one structural study to date has assessed prefrontal subregions, finding reduced gray matter volumes in the left DLPFC and the right OFC in alcoholics with antisocial personalities compared with controls (Laakso et al., 2002). Regarding functional studies, reduced metabolism in the DLPFC has been found in aggressive patients (Hirono et al., 2000) and also aggressive children with epilepsy (Juhasz et al., 2001). In addition, two fMRI studies have observed abnormal DLPFC functioning in APD patients during both an emotional task (Schneider et al., 2000) and also in an inhibition task (Vollm et al., 2004).

**Temporal lobe**

The temporal lobe is the second major brain area traditionally associated with antisocial and aggressive behavior (Raine, 1993). Widespread abnormalities have been reported in several studies. Regarding structure, Kruesi et al. (2004) found a significant association between early-onset conduct disorder (without substance abuse comorbidity) and smaller temporal gray matter volumes. Similar findings of reduced temporal lobe volume have been reported in both incarcerated psychopaths (Dolan et al., 2002) and in APDs (Barkataki et al., 2006).

Regarding function, two SPECT studies have documented reduced temporal lobe functioning in aggressive patients (Volkow and Tancredi, 1987; Amen et al., 1996). Another SPECT study found significant negative correlations between psychopathy and temporal perfusion, particularly left temporal blood flow (Soderstrom et al., 2002). One fMRI study on violent offenders has shown reduced functioning in the temporal cortex compared with non-aggressive controls (Raine et al., 2001). There are also some failures to observe temporal lobe functional impairments; one PET study of violent offenders failed to observe reduced temporal lobe glucose metabolism (Raine et al., 1997a).

These findings are beginning to confirm global structural and functional temporal lobe abnormalities, although results are not entirely consistent. As now outlined, some studies have attempted to localize abnormalities to the middle, anterior inferior, superior and medial temporal (amygdala and hippocampus) regions.

**Middle temporal gyrus.** One PET study found reduced bilateral metabolism in the middle temporal gyrus in aggressive children with temporal lobe epilepsy (Juhasz et al., 2001). A recent SPECT study also revealed reduced blood flow in the right middle temporal gyrus in APD patients (Goethals et al., 2005).
Anterior inferior temporal cortex. Using PET, Wong et al. (1997a) found reduced metabolism in the anterior inferior temporal cortex in two separate groups of violent patients. Similar findings were reported in a SPECT study suggesting a significant reduction in regional blood flow in the left anterior inferior temporal cortex (Hirono et al., 2000).

Superior temporal gyrus. One study has revealed activation deficits in antisocial and psychopathic individuals localized to the right posterior superior temporal gyrus in a semantic processing task (Kiehl et al., 2004). Conversely, a second study observed reduced activation in the left hemisphere superior temporal gyrus in a memory task (Kumari et al., 2006).

Medial temporal cortex. In functional terms, two PET and one SPECT studies have found abnormal glucose metabolism in the medial temporal cortex. Volkow et al. (1995) found significantly reduced metabolism in the medial temporal cortex in psychiatric patients with a history of repetitive violent behavior. The other study showed a similar metabolic reduction in the medial temporal lobes in violent offenders (Seidenwurm et al., 1997). Using SPECT, Soderstrom et al. (2000) found significant blood flow reductions in the right medial temporal region in impulsive violent criminals.

Amygdala. Three structural imaging studies have found significant reductions in the volume of the amygdala in violent offenders (Wong et al., 1997b; Tiitinen et al., 2000) and psychopathic individuals (Yang et al., 2006b). Conversely, no volumetric difference have been found in aggressive patients with temporal lobe epilepsy (Trimble and Van Elst et al., 1999; Van Elst et al., 2000), although the latter study also reported a significantly higher rate of amygdala atrophy (20%) in aggressive patients.

In terms of functioning, abnormalities in the amygdala have been found in one PET study one magnetic resonance spectroscopy (MRS) study, and several fMRI studies. In the PET study, Raine et al. (1997a) found abnormal functional asymmetries in murderers, showing lower left and increased right amygdala functioning, Critchley et al. (2000) using MRS found violent patients with mild mental retardation to have reduced metabolism in the right amygdalo-hippocampal complex compared with non-violent controls. Somewhat surprisingly, two fMRI studies have reported increased amygdala activation in antisocial individuals while viewing negative visual content (Muller et al., 2003) and during an aversive conditioning task (Schneider et al., 2000). In contrast, reduced activation in the amygdala during the processing of affective stimuli has been found in criminal psychopaths (Kiehl et al., 2001), psychopathic individuals (Veit et al., 2002; Birbaumer et al., 2005), and adolescents with conduct disorders (Sterzer et al., 2005).

Hippocampus. Several imaging studies have examined the functional integrity of this region and found abnormalities in murderers (Raine et al., 1998), criminal psychopaths (Kiehl et al., 2001) and violent offenders (Soderstrom et al., 2000; Critchley et al., 2000). With regard to structure, volumetric reductions in the hippocampus have been found in several aMRI studies of psychopaths (Raine et al., 2004), violent offenders with APD (Laakso et al., 2000), violent psychiatric patients (Kumari et al., 2006; Barkataki et al., 2006) and antisocial alcoholics (Laakso et al., 2001).

Other Brain Areas

Several structural and functional imaging studies have suggested that the parietal lobe (particularly the angular gyrus) and anterior/posterior cingulate gyrus may also be compromised in antisocial groups.

Parietal Lobe. Regarding function, reduced metabolism has been found in the superior parietal cortex in aggressive patients (Hirono et al., 2000), murderers (Raine et al., 1997a) and individuals with impulsive personality disorders (Siever et al., 1999). However, no structural MRI study has been conducted to date showing structural impairments in individuals with APD.

Only two studies to our knowledge have assessed functioning in the angular gyrus in individuals with APD. Both found significantly reduced activation, but in different hemispheres. Raine et al. (1997a) in a PET study found that murderers have reduced glucose metabolism in the left angular gyrus. Using SPECT, Soderstrom et al. (2000) found a significant blood flow reduction in the right angular gyrus in impulsive violent criminals.

Anterior and Posterior Cingulate. Several fMRI studies have provided evidence showing functional impairments in the anterior and posterior cingulate cortex (PCC) in APDs. Five fMRI studies have shown reduced activation in the anterior cingulate cortex (ACC) in criminal psychopaths during an affective memory task (Kiehl et al., 2001), APD patients during a working memory task (Kumari et al., 2006), psychopaths during a fear-conditioning task (Birbaumer et al., 2005), and conduct disorder patients during viewing of negative emotional pictures (Sterzer et al., 2005). In addition, an abnormal activation pattern was reported in APD patients during a response inhibition task (Vollm et al., 2004). With regard to the PCC, two functional imaging studies have also found reduced activation in this region in criminal psychopaths (Kiehl et al., 2001) and aggressive patients (New et al., 2002).

Summary and interpretation of imaging findings in antisocial populations

A fairly sizable imaging literature has now been built up on functional and (to a lesser extent) structural brain abnormalities in diverse antisocial groups, although for some structures such as the amygdala and hippocampus, the precise pattern of results is somewhat complex. Taken together, the best-replicated brain imaging abnormality...
found to date across a wide variety of antisocial groups, across structure and function and across different imaging methodologies is the PFC. While dysfunction in the OFC may be associated with poor inhibitory control, emotional decision-making and reward/punishment processing in antisocial individuals, the additional involvement of DLPFC dysfunction may also predispose to response perseveration (life-long ‘revolving door’ antisocial behavior despite repeated punishment), and poor planning/organization (resulting in an occupationally and socially dysfunctional lifestyle).

In addition to the PFC, there is increasing evidence for structure/function impairments in the amygdala (poor fear conditioning), hippocampus (emotion regulation and contextual fear conditioning), temporal cortex (language and memory), anterior cingulate (autonomic functions and emotion regulation) and initial evidence for reduced functioning in the angular gyrus (reading and arithmetic—potentially predisposing to school/occupational failure). Impairments to varying degrees have been documented in antisocial populations in all these neurocognitive processes. It is clear that multiple brain regions may contribute to antisocial behavior for different reasons. Furthermore, future research needs to move beyond the simple identification of single structures impaired in antisocials to the more complex delineation of the neural circuits that are impaired. In all likelihood, no single brain structure is critical in the development of antisocial behavior. Instead, the greater the number of neural impairments across different cognitive and affective domains related to an antisocial lifestyle, the higher the likelihood of an antisocial outcome. The remainder of this article outlines how impairments to the circuitry underlying morality may be one such important additional risk factor for the antisocial spectrum of disorders.

NEUROLOGICAL STUDIES OF ANTISOCIAL BEHAVIOR

Neurological studies of patients suffering trauma to discrete brain regions have provided invaluable data on neural mechanisms predisposing to antisocial behavior. An inevitable limitation of imaging findings is that they are correlational, while, in contrast, the natural accidents that comprise neurological findings bear more directly on causality. Intriguingly, findings from these studies converge with evidence from brain imaging studies, particularly with respect to the PFC.

Two key prefrontal neurological syndromes that bear on antisocial behavior have been delineated on the basis of the timing of neurological damage, and give rise to slightly different antisocial outcomes. ‘Acquired sociopathy’, in which accidental damage occurs to the ventromedial PFC in adulthood, has been shown to result in pseudo-psychopathic, disinhibited, antisocial behavior (Damasio et al., 1990; Damasio, 1994), together with bad decision-making and reduced anticipatory skin conductance responses to stimuli predicting negative outcomes. Nevertheless, with some exceptions (e.g. Blair and Cipolotti, 2000) these patients do not generally exhibit impairments in moral judgment. For example, patient E.V.R. suffered damage to the ventromedial PFC at the age of 35 years, and consequently manifested significant antisocial behavior and reduced skin conductance responsivity to socially meaningful stimuli (Damasio et al., 1990), yet was unimpaired on a moral reasoning task and showed good abstract social insight (Saver and Damasio, 1991).

In the second syndrome of developmental sociopathy, prefrontal damage early in childhood is similarly associated with a history of antisocial and aggressive behavior together with reduced anticipatory skin conductance responses to stimuli associated with punishment that extends throughout childhood and into adulthood. Unlike acquired sociopathy, however, developmental sociopathy is associated with significant impairments in moral reasoning and judgment. For example, two case studies reported by Anderson et al., (1999) show that lesions to either the polar or ventromedial PFC occurring in the first 16 months of life result in lifelong psychopathic-like antisocial behavior, and also impaired social and moral reasoning as assessed by Kohlberg’s moral judgment task. These studies would suggest, therefore, that the PFC, particularly polar and ventromedial sectors, are critically involved in the development (but not necessarily retention) of moral reasoning.

Brain imaging and neurological findings on antisocial behavior show convergence in other ways. The somatic marker hypothesis argues that somatic states associated with previous stimuli/events provide an effective, automatic mechanism to facilitate choice of response options (Damasio et al., 1990; Damasio, 1994). It is further argued that damage to the ventromedial PFC effectively deactivates this mechanism, resulting in poor decision-making and antisocial behavior (Damasio et al., 1990; Bechara et al., 1997). These findings from neurological patients gain some support from the imaging literature. Individuals with APD have been found by MRI to have both a reduced volume of the PFC together with reduced skin conductance during a social stressor that elicits somatic marker states of embarrassment, shame and guilt (Raine et al., 2000). Furthermore, the antisocials with the greatest prefrontal volume reduction show the greatest autonomic impairment. Different clinical neuroscience paradigms consequently converge on a common conclusion that prefrontal cortical impairment is a significant risk factor for the development of antisocial behavior.

NEURAL BASIS TO MORAL JUDGMENTS

Behavior that breaks the moral guidelines set down by society is a fundamental feature of antisocial disorders, and almost defines criminal behavior. In this context,
recent imaging research is beginning to identify which cortical areas are activated when subjects perform tasks involving moral conundrums. This section provides a brief overview of what we have learnt so far about brain mechanisms subserving moral decision-making from 12 fMRI studies. As will be seen, a wide variety of tasks have been used asking somewhat different questions on the neural correlates of morality. Despite this diversity, a number of key brain areas appear to be a common denominator for moral information-processing.

**Regions consistently activated—medial PFC, angular gyrus and ventral PFC**

The ground-breaking study in this area (Greene et al., 2001) focused on the specific difference between making judgments (i.e. ‘appropriate’ or ‘inappropriate’) on ‘moral personal’ dilemmas (e.g. throwing a person out of a sinking lifeboat to save others), and ‘moral impersonal’ dilemmas (e.g. keeping money found in a lost wallet). Moral dilemmas involving a personal component, compared with both impersonal moral dilemmas and nonmoral dilemmas, activated the medial frontal gyrus (BA 9 and 10), the posterior cingulate (BA 31) and both left and right angular gyri (BA 39). This initial study suggested that these structures play a central role in the emotional processes that influence personal moral decision-making.

Studies since 2001 have confirmed the importance of the medial PFC, angular gyrus and posterior cingulate in processing moral stimuli, and at the same time implicated a further structure—the ventral PFC. The medial PFC (and the frontal pole, BA 10) is activated by passive viewing of pictures depicting moral vs nonmoral violations (Moll et al., 2002a; Harenski and Hamann, 2006), making judgments on auditory moral vs nonmoral sentences (Oliveira-Souza and Moll, 2000), passive viewing of morally disgusting vs nonmorally disgusting statements (Moll et al., 2005); difficult vs easy moral dilemmas (Greene et al., 2004), moral decision-making vs semantic decision-making (Heekeren et al., 2003), personal vs impersonal moral dilemmas (Green et al., 2004), judgment on moral vs nonmoral actions (Borg et al., 2006) and ‘utilitarian’ moral decision-making (e.g. acceptance of loss of life for the greater good) vs ‘nonutilitarian’ decision-making (prohibiting a loss of life even though more lives could be saved—Greene et al., 2004). Overall, with only a few exceptions, studies consistently observe activation of fronto-polar cortex (particularly the medial aspects of BA 10) in moral judgment tasks.

The angular gyrus (BA 39—in some studies labeled as posterior superior temporal gyrus) lying in the parietal lobe at the intersection of temporal, parietal and occipital cortices, surrounds the tip of the posterior superior temporal sulcus and is a second region that shows robust activation across different moral judgment tasks. Angular gyral activation initially observed by Greene et al., (2001) has also been observed in passive viewing of pictures depicting moral vs nonmoral violations (Moll et al., 2002a; Harenski and Hamann 2006), making judgments on auditory moral vs nonmoral sentences (Oliveira et al., 2000), moral vs semantic decision-making (Heekeren et al., 2003; replicated in Heekeren et al., 2005), personal vs impersonal moral dilemmas (Greene et al., 2004), judgments on moral vs nonmoral actions (Borg et al., 2006) and difficult vs easy moral judgments (Greene et al., 2004). In addition, Moll et al., (2002b) found activation in the area of the superior temporal sulcus bordering the angular gyrus when responding to unpleasant moral vs unpleasant non-moral statements. The weight of support for angular gyral involvement in moral tasks, in association with the posterior superior temporal sulcus, is consequently as strong as that for medial PFC involvement.

One additional brain region that could not be imaged in the initial study by Greene et al., (2001) due to susceptibility artifact, yet which is being increasingly implicated in moral judgment tasks, is the ventral PFC. This region encompasses the OFC and the gyrus rectus (also broadly termed ventrolateral and ventromedial, respectively). Activation in this region has been found during passive viewing of pictures depicting moral vs non-moral violations (Moll et al., 2002a—right orbitofrontal), responding to unpleasant moral vs unpleasant nonmoral statements (Moll et al., 2002b—gyrus rectus and orbitofrontal), passive viewing of morally disgusting vs non-morally disgusting statements (Moll et al., 2005—orbitofrontal bilaterally), judgments on moral vs non-moral actions (Borg et al., 2006—orbitofrontal), ‘automatic’ moral judgment to high vs low immoral stimuli (Luo et al., 2006—medial orbitofrontal) and moral vs semantic decision-making (Heekeren et al., 2003—ventromedial; replicated in Heekeren et al., 2005).

**Regions less consistently activated—posterior cingulate, amygdala and temporal pole**

While the polar/medial PFC, ventral PFC and angular gyrus encapsulate areas with the strongest evidence for activation during moral tasks, three other regions also need to be considered. Activation of the posterior cingulate has been observed during moral personal vs impersonal dilemmas (Greene et al., 2001), moral vs semantic decision-making (Heekeren et al., 2005) and passive viewing of pictures depicting moral vs non-moral violations (Moll et al., 2002a; Harenski and Hamann 2006). In addition, Green et al., (2004) observed consistent activation of the posterior cingulate across all three of their experimental moral decision-making conditions/comparisons (personal vs impersonal, difficult vs easy and utilitarian vs non utilitarian). Several studies have also reported amygdala activation (Moll et al., 2002a; Greene et al., 2004; Berthoz et al., 2006; Harenski and Hamann 2006; Luo et al., 2006). The temporal pole has also been activated in some
The regulation of moral emotions

The above studies provide important clues as to the neural circuit underlying moral emotions and cognitions. A different question that can be asked concerns which brain areas are involved in the regulation of moral emotions. In addressing this issue, Harenski and Hamann (2006) not only assessed activation during unpleasant moral vs unpleasant nonmoral social picture viewing, but also requested participants to suppress their emotional reactions to both sets of stimuli. As expected, the two brain areas activated by moral (vs nonmoral) conditions were the angular gyrus and the posterior cingulate. In contrast, suppression of emotion to pictures of moral transgressions leads to a reduction in amygdala activation and an increase in polar/medial PFC (area 10) activation. These findings suggest a unique activation role for area 10 of the PFC in both the process of moral decision-making and in the regulation of moral emotions, while deactivation of the amygdala during down-regulation of moral emotions provides further credence to its active involvement in the generation of moral emotions.

MORAL CIRCUITRY AND ANTISOCIAL BEHAVIOR IN AN EVOLUTIONARY CONTEXT

It is becoming increasingly clear that moral judgments are not the sole product of thoughtful introspection conducted in isolation of emotion. In contrast, the imaging literature would suggest that brain systems underlying affective as well as cognitive processes are active when individuals weigh their actions within a moral context. Furthermore, there is likely a complex interplay between cognitions and emotions during the formulation of a moral decision, blurring the distinction between thinking and feeling, at least in modern-day man.

It is likely that some (if not many) of our moral values are shrouded in a deep evolutionary history where emotions—not cognitions—constituted the driving force of moral action. The astonishing success of early hominids was heavily predicated on reciprocal altruism and a resource-sharing social structure. Nevertheless, selfishness (taking but not giving resources) can constitute a competing evolutionary stable strategy that has to be kept in check for the survival of the species. Negative moral emotions likely evolved to counteract the breaking of social conventions. Moral feelings of indignation, disdain, disgust and contempt can give rise to the stronger emotions of outrage and vengeance that then give rise to ostracization of the cheat from the social group, injury or even death. At this level, morality is largely emotion-driven, relatively automatic, and has little or no higher cognitive control component in early hominids. As hominid society became more complex, higher-order cognitive processes likely became increasingly important for both dealing with more complex moral dilemmas, and for regulating the expression of moral emotions.

Despite the evolution of social mechanisms to deter antisocial ‘cheating’ behavior, it has been argued that antisocial behavior is an evolutionary stable strategy—a pre-programmed behavioral approach that maximizes reproductive fitness. Psychopathy has been viewed as the full expression of this ‘cheating’ strategy (Raine, 1993). At low base rates within the population, psychopaths can be successful in extracting resources from other individuals before moving on to other social groups to avoid the consequences of moralistic rage and retributive justice. Psychopathic traits of superficial charm, egocentrism, manipulation, pathological lying and deception, promiscuous sexual behavior, lack of remorse and guilt, superficial relationships and general parasitic lifestyle are viewed as key components of this evolutionary stable strategy that are observable at a surface level (Raine, 1993). Importantly, psychopaths adopt a transient, unstable, ‘stimulation-seeking’ lifestyle, moving from one place to another to avoid ultimate detection—an essential antidote to the moral medicine that would otherwise be metered out to them.

In this evolutionary context, an essential component of this successful cheating strategy must be a gene machine that lacks a core moral sense. One way to create such a cheating machine would be to engineer individuals lacking the neural circuitry essential for moral feelings and behavior. One prediction generated by this model is that antisocial and psychopathic individuals would manifest impairments in the brain mechanism requisite for morals, particularly those neural processes critical to the
experiencing of moral emotions—the brakes on rule-breaking behavior.

**NEURAL MODELS OF MORALITY AND ANTISOCIALITY**

The review outlined above concludes that the brain regions most implicated to date in antisocial, aggressive and psychopathic behavior are the prefrontal (dorsolateral and ventral) and temporal cortices. Some evidence also exists implicating the amygdala, hippocampus, angular gyrus and anterior cingulate. It has also been argued that the best-replicated neural correlates of morality are the polar/medial PFC, ventral PFC and angular gyrus, with significant support also for the posterior cingulate and amygdala. We also argue that a critical driving force of antisocial and psychopathic behavior is a disruption to the neural circuitry underlying moral thinking and feeling. Figure 1 juxtaposes these two sets of empirical data to create an initial neural model of morality and antisociality, highlighting those regions impaired only in antisocial populations (red), areas activated only by moral judgment tasks (green) and areas common to both antisociality and morality (yellow).

It can be seen both that there are substantial areas of overlap between antisocial/psychopathic behavior and moral judgment-emotion, and also significant differences. Brain regions common to both include ventral and polar/medial PFC sectors, the amygdala and angular gyrus/posterior superior temporal gyrus. A key difference is that while there is increasing evidence for hippocampal and anterior cingulate impairment in antisocial/psychopathic behavior, cognitive components of morality studies consistently fail to observe selective activation in these structures. Similarly, there is some replicable evidence implicating the posterior cingulate in moral judgment tasks, but evidence implicating this region in antisocial behavior is to date sparse. A further qualification is that only two studies have documented reduced functioning in the angular gyrus in antisocial populations, although few studies have explored this parietal structure which should be a target region-of-interest in future imaging studies.

The partial overlap of structures implicated in antisocial populations and moral judgment tasks gives rise to the hypothesis that some of the brain impairments found in antisocial individuals disrupt moral emotion/decision-making, which in turn predisposes the individual to rule-breaking, antisocial behavior. A critical question raised by this theoretical perspective of antisocial behavior concerns which component of ‘morality’ is impaired—cognition (i.e. moral ‘reasoning’), emotion (i.e. moral ‘feelings’) or both?

A tentative answer would be that the emotional component is most impaired in antisocial, psychopathic populations. Regarding basic cognitive processes involved in moral decision-making, at a fundamental level there is little question that almost all criminal and psychopathic individuals know right from wrong. While some evidence exists for a difference in level of moral reasoning in delinquent, criminal and psychopathic groups (Raine, 1993; Blair, 1995), antisocial behavior could cause differences in moral thinking, rather than vice versa. That is, living an antisocial way of life may change moral thinking to justify the individual’s repeated antisocial actions and reduce cognitive dissonance. Furthermore, it has also been argued that psychopaths show excellent (not poor) moral reasoning ability when discussing hypothetical situations—their real failure comes in applying their excellent moral conceptual formulations to guiding their own behavior (Cleckley, 1976).

**The feeling of what’s right and wrong**

Given the above, it is suggested that it is predominantly the feeling of what is moral that is deficient in antisocial groups, rather than the knowing of what is moral. This moral feeling, centered on the PFC and amygdala, is the engine that translates the cognitive recognition that an act is immoral into behavioral inhibition—and it is this engine that functions less well in antisocial, violent and psychopathic individuals.

We also hypothesize that while deficits in the affective component of morality is the primary impairment in antisocial individuals, cognitive components of morality could also be compromised for three reasons. First, both the polar/medial PFC and posterior cingulate have been shown to play an important role in self-appraisal and self-reflection (Ochsner et al., 2005; Johnson et al., 2006), while the OFC is also involved in the process of self-perception and insight (Beer et al., 2006). If an individual is unable to relate back onto themselves the negative emotion associated with the thought of perpetrating an immoral act due to impairments to either the medial PFC or posterior cingulate, they may become predisposed to rule-breaking behavior, despite intact emotional processing. Second, cognition and emotion in moral decision-making cannot be easily dissociated. The medial PFC is activated during the suppression of moral emotions (Harenski and Hamann 2006). If an individual lacks this regulatory control, there may be an inability to down-regulate moral outrage triggered by a third party’s negative acts, resulting in impulsive, reactive aggression by that individual, and/or more planned, controlled retaliatory actions. Third, and relatedly, the angular gyrus has been speculated to be associated with a sense of responsibility for one’s actions (Borg et al., 2006); while intact moral emotions may normally place a brake on rule-breaking behavior, the lack of a sense of responsibility may move the individual more in an immoral direction if the rewards are sufficient. Consequently, both emotional and cognitive components of the morality
circuit may be implicated in antisocial, psychopathic behavior.

CONCLUSIONS AND SUMMARY
In summary, we have argued the following:

(i) brain regions compromised in antisocial, violent and psychopathic populations include both dorsal and ventral regions of the PFC, amygdala, hippocampus, angular gyrus, anterior cingulate and temporal cortex including the superior temporal gyrus;

(ii) regions activated during moral decision-making in normals include the polar/medial PFC, ventral PFC, angular gyrus, amygdala and posterior cingulate;

(iii) brain areas associated with both moral reasoning and antisocial behavior significantly overlap;

(iv) the rule-breaking, immoral behavior of antisocial and psychopathic individuals may in part be due to impairments in those brain regions subserving moral cognition and emotion;

(v) while impairments to the moral emotional system may be primary in antisocials, disruption of moral cognitive and cognitive-emotional systems are also possible.

Fig. 1 A schematic diagram of brain regions impaired only in antisocial groups (red), activated only in moral decision-making (green) and regions common to both antisocial behavior and moral decision-making (yellow).
This neuro-moral theory of antisocial, violent and psychopathic behavior must be regarded as provisional. The precise foci of the likely multiple neural deficits in antisocial groups (particularly within the large regions compromising prefrontal and temporal cortex) remains to be delineated, while an understanding of the neural basis to moral decision-making is clearly still in its infancy. For example, there are hints that temporal pole impairment/activity may be implicated in both antisociality/morality, respectively, yet, limitations in the evidence to date precluded inclusion in Figure 1. A compromised moral circuit will be only one of multiple etiological processes ultimately found to predispose to the complex constructs of antisocial/aggressive/psychopathic behavior. Interactions between neural and social risk factors for antisocial behavior cannot be ignored (Raine et al., 1997b; Caspi et al., 2002), raising yet another layer of complexity to a full elucidation of antisocial disorders.

Despite these caveats, we believe that a neural moral hypothesis of antisocial disorders is worthy of examination in future studies. Neuroscientists and lawyers alike are beginning to raise important questions about the implications of new neuroscience knowledge for society, the law, and civil liberties (Morse, 2004), leading to the beginning of a new sub-discipline of 'neuroethics' (Farah, 2004). Psychopaths may not be 'morally insane' in any strict legal sense as they are cognitively capable of distinguishing right from wrong, but if they lack the capacity for the feeling of what is moral due to neurobiological impairments beyond their control, are they fully responsible for their criminal behavior? If not, what are the implications for punishment and our concepts of both justice and retribution? This challenging question that lies at the interface of law, neuroscience and neuroethics, begs for further enlightenment from future systematic imaging research on both morality-processing and antisocial behavior.

Conflict of Interest
None declared.

REFERENCES


