OPINION

Moral cognition and its neural constituents

William D. Casebeer

Identifying the neural mechanisms of moral cognition is especially difficult. In part, this is because moral cognition taps multiple cognitive sub-processes, being a highly distributed, whole-brain affair. The assumptions required to make progress in identifying the neural constituents of moral cognition might simplify morally salient stimuli to the point that they no longer activate the requisite neural architectures, but the right experiments can overcome this difficulty. The current evidence allows us to draw a tentative conclusion: the moral psychology required by virtue theory is the most neurobiologically plausible.

Good moral reasoning is extremely important for *Homo sapiens*. Our lives are more fruitful if we recognize salient ethical norms and reason effectively about their application to our own situations. We are social creatures, and if we are to flourish in our social environments, we must learn how to reason well about what we should do. Despite its importance for our proper functioning, until recently the neural mechanisms of moral cognition were not well studied. This is unfortunate, as co-evolution between the neural constituents of moral cognition and the moral psychologies that are required by the main ethical theories is necessary if we are to make progress in understanding how effective ethical reasoning is embodied in the brain. To make such progress requires us to probe the nature of moral judgement and its relationship to the experimental regimens that are used to explore such constituents. Here, I briefly review the neural mechanisms of moral cognition, discuss methodological pitfalls and consider issues that might inform future experimental work. Ultimately, the current situation makes the moral psychology that is required by virtue theory the most neurobiologically plausible, although this is a tentative, defeasible conclusion, and more work is needed to confirm it.

Moral theories and moral cognition

To study the neural mechanisms of moral cognition, one must delimit the field of inquiry. What does ‘moral cognition’ encompass? This depends on how we construe the domain of moral theory. Although all moral theories claim to speak to what an agent should do (this is what makes them distinctively moral), they disagree about the substance of such recommendations and the moral psychologies that are required for effective reasoning and action. The three main classic moral theories in the Western tradition are utilitarianism, deontology and virtue theory.

The typical utilitarian, such as the British philosopher John Stuart Mill (1806–1873), thinks that one should take that action (or follow that ‘rule’) that, if taken (or followed), would produce the greatest amount of happiness for the largest number of sentient beings, where happiness is the presence of pleasure or the absence of pain (and where pleasure and pain are given more sophisticated readings than mere affective satisfaction). The second flavour of utility, ‘rule utilitarianism’, is probably the most popular[1].

Deontologists, exemplified by the Prussian philosopher Immanuel Kant (1724–1804), do not emphasize the consequences of actions, as utilitarians do. Instead, they focus on the maxim of the action — the intent-based principle that plays itself out in an agent’s mind. We must do our duty, as derived from the dictates of pure reason and the ‘categorical imperative’, for duty’s sake alone. Deontologists are particularly concerned to highlight the duties that are owed to each other by free and reasonable creatures (paradigmatically, humans). Maximizing happiness is not the goal; instead, ensuring that we do not violate another’s rights is paramount[2].

Virtue theorists, such as the Greek philosophers Plato (427–347 BC) and Aristotle (384–322 BC), make paramount the concept of ‘human flourishing’[3,4]; to be maximally moral is to function as well as one can given one’s nature. This involves the cultivation of virtues (such as wisdom) and the avoidance of vices (such as intemperance), and is a practical affair.

Each approach asks different things of us cognitively. What follows is an abbreviated discussion of each theory’s moral psychology. To make the appropriate judgement about what one should do, the utilitarian would, at least in morally problematic cases, require that a moral agent could recognize and compute salient utility functions. We would then be moved to act on such judgements by cultivation of appropriate altruistic fellow-feeling or, in many cases, merely by self-concern (as utility will often be maximized by having each of us focus on our own happiness as well as the happiness of others). So, in terms of raw computations, a utilitarian moral psychology would require some mechanism for learning what actions or rules would eventually produce happiness. Either implicitly or explicitly, utilitarian computations would constitute the bulk of our moral cognitive capacity. Whether we act on the outcome of those judgements might require some derivative character development (such as the cultivation of concern for the happiness of

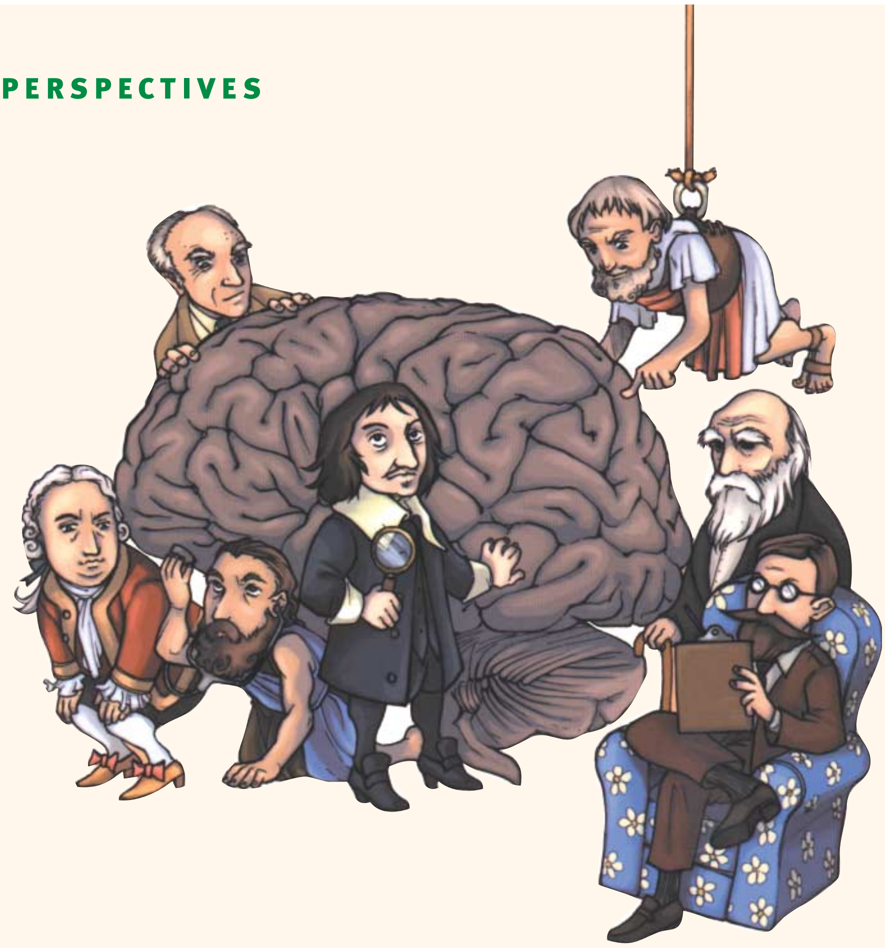

Figure 1 | **A carnival of philosophers.** Kant is shown at lower left, Aristotle at upper right. Reproduced, with permission, from REF. 34 © (2002) MIT Press.

others), and this would require appropriate training of the emotions.

A Kantian moral psychology would be different. The ability to 'reason purely' about the demands of the categorical imperative (the heuristic that is used by Kant to capture our respect for those things that make morality possible — autonomy and rationality) would be the most important part of our cognitive equipment. The best known formulation of the categorical imperative requires that we act only on maxims that we can will to become a universal law; other maxims are morally impermissible. For example, you could not universalize the maxim that allows you to lie to achieve some end; such a maxim requires that others act with a different intention, of delivering and receiving only true utterances (or else the lie would not be effective). The maxim cannot be made universal; this conceptual truth does not require experimentation to confirm it. We therefore have a perfect (that is, exceptionless) duty not to form the intention to lie. Of note, Kant requires that we be moved to do our duty by the demands of duty alone; if something else (say, the desire to be liked) is moving us, our action is not morally praiseworthy as it plays more to our animal nature than our rational (and so human) nature. Kant is thought to give short shrift to character development and related issues, although recent work has 'softened up' this position[5]. What exactly the cognitive capacity to reason purely in the Kantian manner would look like has not been a subject of extensive investigation; it would require at least the ability to check universalized maxims for logical consistency in a manner that is separable from the taint of affect and emotion.

Finally, virtue-theoretic moral psychology is often thought to be the richest of the three. A virtuous person must be able to reason well about what states of being would be most conducive to the best life. What type of person must I become if I am to experience eudaimonia (variously translated from the Greek as flourishing, proper functioning or happiness)? To act on the outcomes of my judgements, I must train my character so that my appetites and 'spirit' are coordinated smoothly with the demands of good reason. Virtuous people are moved to do the appropriate thing at the appropriate time; they become angry at unjust events, are sympathetic to recipients of wrong-doing, and so on. Virtue theorists focus on the appropriate coordination of properly functioning cognitive sub-entities. Moral reasoning and action are therefore 'whole-psychology, whole-brain' affairs. Jokingly, then, it could be said that these approaches emphasize different brain regions: frontal (Kant); pre-frontal, limbic and sensory (Mill); and the properly coordinated action of all (Aristotle) (FIG. 1).

Kant would say that moral reasoning is a robustly rational affair, where 'rational' is given a strict interpretation. With Aristotle, however, I think it is more useful to treat moral judgement in a deflationary manner. Given that the domain of what constitutes a moral judgement is itself in contention, we would be best served by casting our nets widely, narrowing them appropriately as the neurobiological, psychological and normative aspects of morality co-evolve (admittedly, casting our net so widely might bias us initially towards a virtue-theoretic moral psychology; the give and take required by co-evolution of theories at all three levels of analysis would, hopefully, correct any such bias eventually). As a first cut, then, moral cognition comprises any cognitive act that is related to helping us ascertain and act on what we should do. Non-human animals (for example, primates and other social animals) might also engage in robust moral reasoning (see, for example, REF. 6).

My push for this deflationary conception of moral judgement is driven by the recognition (but also has as the upshot) that moral cognition might not be a tightly defined 'natural kind' in the sense that other cognitive phenomena might be. For example, the domain of the neural mechanisms of visual cognition, owing to the relatively restricted range of information that is processed by the visual modality, might be more tightly constrained than the domain of the neural mechanisms of basket-weaving. In that sense, the former is a more robust, natural kind than the latter, and is therefore an easier target for neurobiological study. Moral reasoning probably falls somewhere between these two extremes and is still worthy of study by neurobiologists, although this fact might make it more difficult to progress experimentally[7].

Critics might argue that such a co-evolutionary strategy commits the naturalistic fallacy of inferring what should be from what is. The exact nature and status of the naturalistic fallacy is subject to debate (for a summary, see chapter two of REF. 8). Note, however, that the two most famous arguments against naturalism about ethics, Hume's law[9] and G. E. Moore's open question argument[10], do not stand up against some contemporary naturalized ethical theories. Both of these arguments rely on an analytic/synthetic distinction that many philosophers agree collapsed in the twentieth century. In addition, Hume's argument rules out deductive

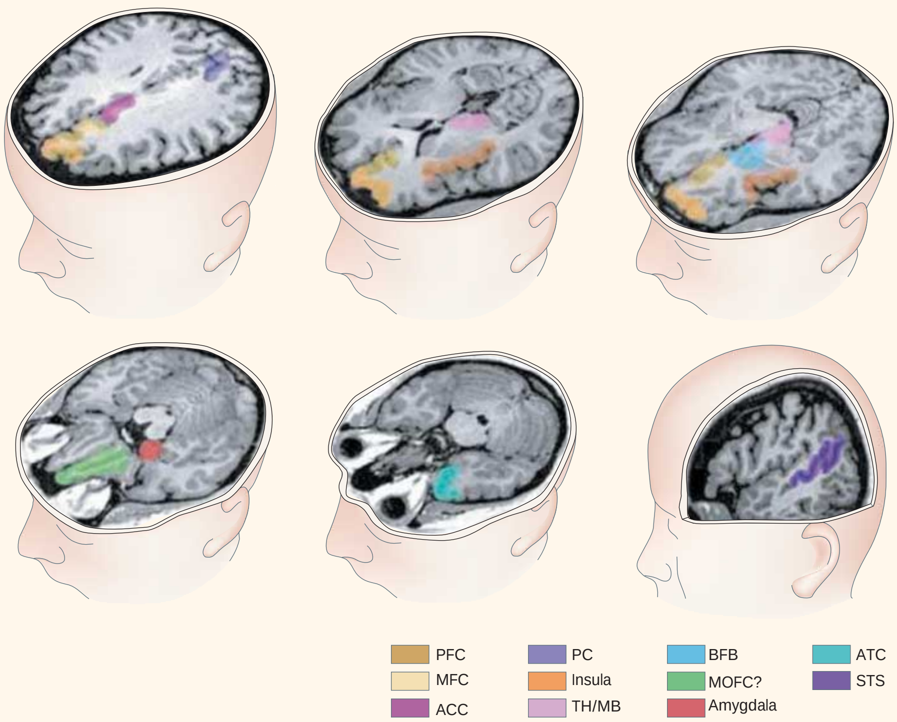

Figure 2 | **Five horizontal slices and one sagittal cross-section of a brain showing many morally relevant neural areas.** ACC, anterior cingulate cortex; ATC, anterior temporal cortex; BFB, basal forebrain; FPC, frontopolar cortex; MFC, medial frontal cortex; MOFC, medial orbitofrontal cortex; PC, precuneous; STS, superior temporal sulcus; TH/MB, thalamus/midbrain. Images courtesy of J. Moll.

relationships between facts and norms, but not necessarily abductive ('inference to the best explanation' style) relationships. And Moore himself admitted, in the second version of his *Principa Ethica*, that his argument best applies to the two forms of naturalized ethics that he attacks in the book: Spencer's evolutionary ethic and hedonism. I agree that both of these naturalized ethics are poor moral theories, but I disagree that Moore has offered an argument that is general to all attempts to naturalize ethics.

Keep in mind that no good naturalized ethical theory will say that all facts are normative facts, nor that all existing states of affairs — merely because they are 'natural' in the sense that they were produced by natural processes — are good. Unlike Greene in his companion piece in this issue[11], I think that the neurobiological facts support a version of relational moral realism, but this discussion is beyond the scope of my paper. Ultimately, even the most ardent anti-naturalist would admit that, at the very least, our moral theories must require us to carry out cognitive acts that are also possible for us to implement. The goal of naturalized ethics is to show that norms are natural, and that they arise from and are justified by purely natural processes. If this can be done, then the naturalistic fallacy is not actually a fallacy (it merely amounts to saying that you don't have a good naturalized ethical theory yet).

Emotion and affect

The rich and diverse literature on the neural mechanisms of moral cognition can be usefully divided into three branches: the moral emotions, theory of mind and abstract moral reasoning. I will briefly discuss the connections between cortical areas and the limbic system that are necessary for good moral judgement, the neural correlates of theory of mind (TOM) and how they manifest themselves in moral judgement, and useful conceptual tools for thinking about abstract moral reasoning.

The moral emotions are crucial for effective moral cognition. They motivate action, serve as markers of value, are vital for coordinating group activity, and help to filter out and highlight certain aspects of the moral calculus. In mammals and reptiles, the brain's regulatory core is situated in the brainstem/limbic axis, and it subserves important activities such as breathing, arousal and the coordination of drives (for food, sex, oxygen and so on) with perceptions (turkey sandwich there, attractive mate here). The basic emotions associated with hunger, thirst, sexual desire and the like are powerful motivators; as our brain's cortical capacity expanded during evolution, 'newer' frontal brain regions remained connected to and were innervated by this regulatory core. No wonder then that these basic, survival-laden emotions serve as the platform on which the moral emotions (and effective moral reasoning) are built.

Exploring the role of the prefrontal regions in connecting limbic areas to frontal areas is difficult; studies of humans with focal brain damage and experimental lesions in monkeys have indicated a relationship between prefrontal cortex (PFC) and planning, decision-making, emotion, attention, memory for spatio-temporal patterns, and recognition of a mismatch between intention and execution. The precise nature of the relationship between these functions and the contribution of the PFC remains unclear, in part because the temporal and spatial resolution of functional magnetic resonance imaging (fMRI) is limited, and because we are still ascertaining the network-level properties of this area.

Despite these difficulties, the link between moral decision-making, social cognition and the emotions is becoming clearer. Converging results from lesion and imaging studies indicate that damage to the ventral and medial PFC is consistently associated with impairments in practical and moral decision-making[12–15]. Patients with focal ventromedial lesions show abnormally flat (emotionless) responses when shown emotional pictures, and perform poorly on tasks where feelings are needed to guide complex self-directed choices.

Electroencephalogram (EEG) studies of children with self-control disorders[16] also support a link between ventromedial PFC and moral emotions, as does fMRI of normal subjects. For example, viewing scenes that evoke moral emotions produces activation in the ventromedial PFC and the superior temporal sulcus[17,18].

Within the ventral PFC, orbitofrontal cortex (OFC) is crucial for cueing morally appropriate behaviour in adulthood and acquiring moral knowledge during childhood; although patients with adult-onset and childhood-onset OFC damage showed similar abnormal socio-moral behaviour, their scores on standardized tests of moral reasoning differed. Those with early damage performed poorly on the tests, exhibiting the egoistic reasoning that is typical of a ten year old, whereas adult-onset subjects performed normally despite their abnormal behaviour[15].

The PFC receives important inputs from both sensory and limbic areas. The limbic system is a highly interconnected set of

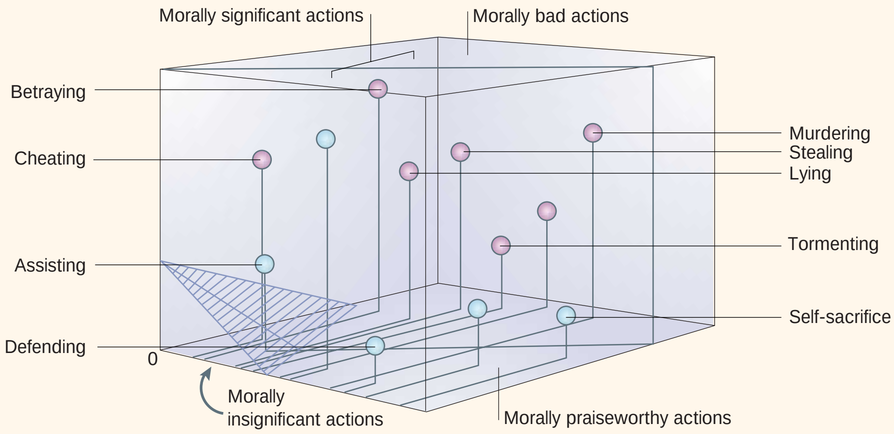

Figure 3 | **A 'moral state-space'.** The axes of this reduced state-space would correspond to functionally salient groups of neurons, and regions of the state space might correspond to moral theories or they might help us to identify undiscovered moral concepts. Modified, with permission, from REF. 35 © (1998) Kluwer Academic Publishers.

subcortical regions (including the hippocampus, amygdala, hypothalamus and basal forebrain) and the cingulate cortex. The activity of this system is modulated by the neurotransmitters dopamine, serotonin, noradrenaline and acetylcholine, and changes in the levels of these substances can greatly affect sex drive, moods, emotions and aggressiveness. The proper operation of the system as a whole is crucial for effective moral judgement. FIGURE 2 shows some of the brain areas that are important in moral cognition.

The amygdala, for example, is part of the complex reward circuitry involving the positive emotions[19]. It is likely that the amygdaloid complex modulates the storage of emotionally important and arousing memories; events that are important to survival provoke specific emotions and, with amygdala activity, are more likely to be permanently stored than neutral events. The amygdala is also crucial for aiding retrieval of socially relevant knowledge about facial appearance[20]; three subjects with total bilateral amygdala damage were asked to judge the trustworthiness of unfamiliar people, and all three judged unfamiliar people to be more approachable and trustworthy than did control subjects[20].

Hippocampal structures are essential for learning and remembering specific events or episodes, although permanent memory storage lies elsewhere in the cortex. The hippocampus, parahippocampal cortex, entorhinal cortex and perirhinal cortex all seem to be important for the processing and retrieval of salient 'me-relevant' memories. In moral judgement, the hippocampus might facilitate conscious recollection of schemas and memories that allow past events to affect current decisions.

The cingulate cortex has a number of subregions with different functions: regulation of selective attention, regulation of motivation, and detection of malcoordinated intention and execution are associated with anterior regions (anterior cingulate cortex, ACC). Rostral ACC activation (along with the nucleus accumbens, the caudate nucleus and ventromedial (VM)/OFC) is needed for cooperative behaviour among subjects playing a version of the 'prisoner's dilemma'; the hyperscanning methodology used to obtain these results is especially promising[21] (see later in text). Other work supports the conclusion that the ACC is crucial for identifying times when the organism needs to be more strongly engaged in controlling its behaviour[22–25].

Theory of mind and moral judgement

This brief exploration of limbic areas and their connections undervalues the role of PFC in the second important research area in moral cognition: social judgement and TOM. Our ability to know what others are thinking so that we can interact fruitfully with them is vital — it underlies our ability to empathize with others, to judge how they might react in response to our actions, and to predict the subjective consequences of our actions for conspecifics. Studies of children with autism indicate that TOM might be subserved by the aggregate neural activity of the OFC, the medial structures of the amygdala and the superior temporal sulcus (STS). The circuit that is formed by the last two structures might mediate direction-of-gaze detection (a crucial component of our ability to infer what others might be thinking about), all three locations are probably involved in mediating shared attention, and the specially coordinated action of all three might therefore constitute TOM processing[26].

Relatedly, 'mirror neurons' in the PFC of the macaque monkey respond either when the monkey makes a specific movement, such as grasping with the index finger and thumb, or when it sees another making the same movement[27]. This indicates that when the animal sees another make the movement, the premotor cortex generates incipient motor commands to match the movement. These signals might be detected as off-line intentions that are used to interpret what is seen (for example, 'the dominant male intends to attack me'). Mirror neurons might therefore bootstrap full-blown TOM into existence through an inner simulation of the behaviour of others. TOM is probably what allows a chimpanzee to know whether a high-ranking male can see the food she is grabbing or whether the food is occluded from view and can be taken without fear[28]. Robust TOM is necessary for healthy moral judgement; it is also associated with our ability to lie, but is nonetheless probably necessary if a whole host of morally important cognitive abilities are to be realized. For an excellent review of TOM mechanisms, see REF. 29.

Abstract moral reasoning

Probably the most difficult aspect of the neural mechanisms of moral cognition is the constituents of abstract moral reasoning. The most important forms of moral reasoning that we rely on daily, involve background social skills, tacit use of TOM, ready-at-hand action patterns and interpretive schema, and the like. Much of our day-to-day moral reasoning does not involve highly convoluted moral modelling; mostly, we can rely on skills and habits of character as informed by conditioned emotion and affect (indeed, Haidt claims that abstract moral reasoning is a completely post-hoc affair and is almost never the direct cause of moral judgements)[30]. Nonetheless, abstract moral reasoning is sometimes necessary. It probably depends on brain structures that subserve morally neutral abstract thought (such as a capacity to model the consequences of an action) and practical reasoning about how to accomplish things. For example, in a classic moral dilemma, such as the trolley problem (in which one has to decide whether to allow an out-of-control trolley to continue down a track where it will strike five people or whether to throw a switch diverting it onto a track where it will strike only one person, explored in detail in fMRI work by Greene *et al.*[31]), higher-order cognitive abilities such as planning, executive flexibility and strategy application (see REF. 32) are needed.

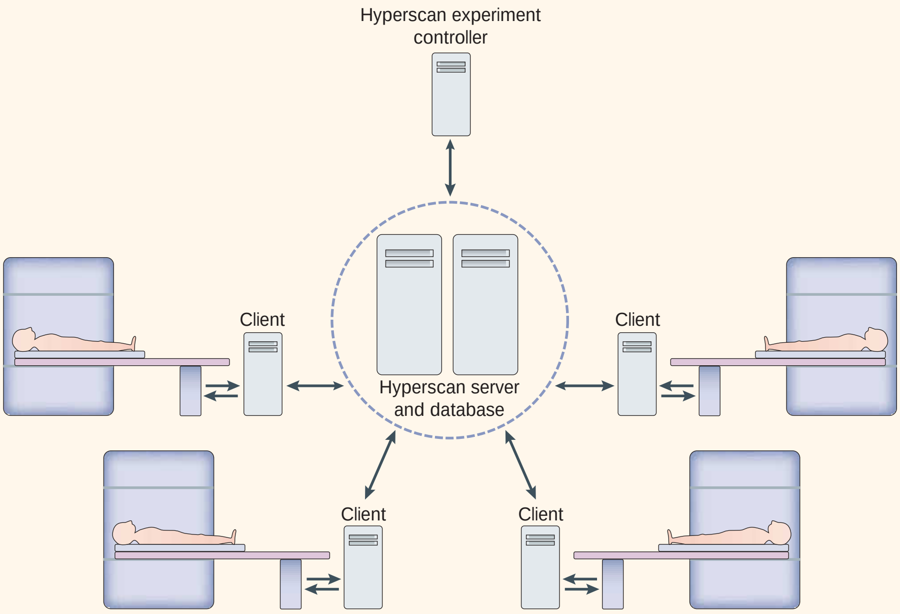

Figure 4 | **A schematic of four subjects engaged in a hyperscanning experiment, interacting through an internet connection.** Such a setup allows simultaneous functional neuroimaging of subjects as they interact with each other. The data collected by each functional magnetic image is passed, through a client, to a hyperscan server and database, which is regulated by an experiment controller. Image courtesy of Read Montague.

These capacities might be realized in cerebral cortex by transient cortical networks which Fuster calls ‘cognits’[33].

The difficulty we have in understanding the neural basis of moral reasoning is indicative of two things: first, that we still need both better theoretical frameworks to understand higher-order cognitive capacities and better network-level tools for probing activity; and second, that such capacities might be overvalued relative to the work that they perform in our cognitive economy (in some respects, then, eliminativism might be called for; see REF. 34).

One device that might be useful for helping us to organize abstract moral reasoning is a moral state-space (a concept first articulated by P. M. Churchland)[35]. We can think of much of the activity of frontal cortex and the limbic/brainstem axis as consisting of a moving point in an n-dimensional space, where n could (in complex cases) be determined by making the activity of every neuron that is involved in the system an axis of that space (in some cases, an axis might be constituted by a single neuron, which might explain results about how single neurons in PFC can seem to encode ‘rules’[36]). Reducing the dimensions of this space enables us to capture its principal components, which might themselves correspond to traditional moral concepts that have been explored by ethicists for the last 2,500 years. The idea of a moral state-space (FIG. 3) allows us to aggregate various cortical regions involved in the processing of moral concepts: if we identify neurons or relevant populations of neurons, and tag each of them as being a dimension of the space, using the right statistical tools (principal or independent components analysis primarily), we can reduce the dimensionality of the space to something that is more manageable.

To behave morally would be to have this state-space allocated appropriately (presumably by the conjunction of experience and the ontogeny of native neurobiological equipment) so that one is maximally moral (which, in the case of virtue theory, means being maximally functional). The axes of this reduced state-space would correspond to functionally salient groups of neurons, and regions of the state-space might correspond to the ‘big three’ traditional moral theories that were discussed earlier, or they might help us to identify undiscovered moral concepts. The idea of a moral state-space is one way to conceptually unify disparate brain activity that is related to moral cognition.

Consilience with virtue theory

The evidence, albeit tentative, that we have discussed lends more credence to the moral psychology that is required by virtue theory. Empirically successful moral cognition on the part of an organism requires the appropriate coordination of multi-modal signals conjoined with appropriately cued executive systems that share rich connections with affective and conative brain structures that draw on conditioned memories and insight into the minds of others, so as to think about and actually behave in a maximally functional manner. There is clear consilience between contemporary neuroethics and Aristotelian moral psychology. A co-evolutionary strategy, then, would suggest that some version of pragmatic Aristotelian virtue theory is most compatible with the neurobiological sciences.

The localization work that is mentioned in this paper uses various techniques and experimental regimens. The stimuli normally range from sentences to small photographs. The constraints of rigorous experimental design mean that tests are sometimes conducted in highly artificial situations; ecological validity for moral reasoning is difficult. Moral cognition exhibits several characteristics that make it difficult to capture in the fMRI chamber[37]. A list of these characterisitics follows.

Moral cognition is ‘hot’. Owing to evolutionary history, affective and conative states are part and parcel of effective moral judgement. This is endemic to moral reasoning (consider the Damasios’ patient E.V.R., whose PFC damage, like that of Phineas Gage, disrupted the connections between limbic ‘somatic markers’ and frontal cortex, resulting in poor moral judgement[14]). Unfortunately, hot cognition is difficult to capture in artificial settings.

Moral cognition is social. Several crucial components of the neural constituents of moral cognition aim to achieve appropriate behaviour in social and group settings. This is no accident: animal and human groups are social groups. Social environments are difficult to simulate in the scanner. A notable methodological improvement in this area is Montague’s use of multi-scanner ‘hyperscanning’ methodology[21], in which several subjects can interact simultaneously while being scanned (FIG. 4). Although the technology is currently used to link-up subjects from multiple states concomitantly, it could be used to monitor the interactions of subjects who are within sight of each other. Minimally, this technology offers added efficiency in studying social interactions; maximally, however, it adds another dimension to the study of the neural mechanisms of social reasoning. As Montague *et al.* note, “studying social interactions by scanning the brain of just one person is analogous to studying synapses while observing either the presynaptic neuron or the postsynaptic neuron, but never both simultaneously … synapses, like socially interacting people, are best understood by simultaneously studying the interacting components”[21] (FIG. 4).

Moral cognition is distributed. Evolution does not build from scratch, but instead tends to work with what is present. Socio-moral behaviour is rooted in the brainstem/limbic axis and PFC, with input and recurrent connections to and from sensory and multimodal cortices and frontal lobe areas: so it involves more-or-less the entire brain. The reduced stimuli conditions that are necessary to do work in an fMRI chamber might not robustly engage our entire suite of neuroethical equipment.

Moral cognition is context-dependent. In one case I might praise you for stealing ("nice work removing that weapon from the terrorist headquarters"), and in another I might condemn you ("please return the candy bar that you stole to its owner"). Experimental setups need to take this context sensitivity into account. Experiments with the 'trolley problem' do a nice job of teasing apart context-sensitive strands of moral judgement[31].

Moral cognition is genuine. Emotion, reason and action are bundled together. Selection forces operate on actual behaviour, not on hypothetical behaviour. Our moral cognitive equipment has evolved to effectively coordinate all aspects of our mind/brain so as to take action that allows us to function properly. Experimental regimens that isolate 'dry' thinking-about-things-moral from 'wet' here-I-am-doing-moral-things can unnecessarily restrict the scope of the neural mechanisms that are activated.

Moral cognition is directed. Moral cognition is about things, broadly construed: how do we interact with the world in a fecund manner? What must I do to function properly? Effective moral cognition is a developmental issue; our socio-moral cognitive system becomes more skilled at navigating a complex physical–social world as time passes. Isolating the act of moral judgement (a knowing 'that') from the idea of knowing how to act in the world can be misleading. When in the scanner, I will push the man onto the railroad track to stop the oncoming train; what I would actually do in the real world is more difficult to predict.

This review, and these general observations, lead to several pieces of methodological advice for researchers studying the neural bases of moral cognition.

Make things explicit. Ensure that explicit consideration is given to the background moral theory that is affecting the research question being answered and the stimuli domain being used to probe that question. When you use the word 'moral' in your research report, ask yourself which moral theory you have in mind, and whether implicit but unexamined background assumptions are causing you to ignore salient data or to choose irrelevant problems.

Adjust experimental regimens accordingly. Confirm that you have a theoretically rich (but nonetheless fallible) foundation for the moral intuitions informing the experiment. This requires reviewing the few big-picture survey articles available in the field[7,29,32,37] (see also REFS 35,38; REFS 39–41 will also be useful). Consider that the term 'moral' is a theoretical term, and so the link between the stimulus regimen, the problem set and the theory being tested is more complex than one might think. For example, to assume that certain stimuli sentences are empty of moral content merely because they are purely 'factual' is to load the dice against moral realism; or, to assume that certain social pictures are morally neutral merely because they aren't threatening is to load the dice against a robustly social conception of morality.

Keep ecological validity in mind. Is there any way that you can make the problem and stimulus set more closely resemble an actual socio-moral problem and its accompanying embedded environmental stimuli? Methodologies that are socially robust and that involve interacting with more than just sentential input will be more likely to meet this requirement.

Doing all of this intelligently is difficult. There are several problems on the frontiers of brain science, and tackling the neural constituents of moral cognition is surely in the 'top ten' in terms of both difficulty and importance. I applaud the researchers who are accomplishing this ground-breaking work. Only by doing it can we seek consilience between normative moral theory, moral psychology and moral neurobiology; and only by doing these things can we hope to improve our ability to develop and instil good moral judgement in ourselves.

*William D. Casebeer (Major, USAF) is at the
US Air Force Academy, HQ USAFA/DFPY,
2354 Fairchild Drive, Suite 1A10,
US Air Force Academy,
Colorado 80840, USA.
e-mail: William.casebeer@usafa.af.mil.*

doi:10.1038/nrn1223

1. Hooker, B. *Ideal Code, Real World: A Rule–Consequentialist Theory of Morality* (Oxford Univ. Press, New York, 2000).
2. Hill, T. *Human Welfare and Moral Worth: Kantian Perspectives* (Oxford Univ. Press, New York, 2002).
3. Broadie, S. *Ethics With Aristotle* (Oxford Univ. Press, New York, 1991).
4. Hursthouse, R. *On Virtue Ethics* (Oxford Univ. Press, New York, 1999).
5. Munzel, G. F. *Kant's Conception of Moral Character: The 'Critical' Link of Morality, Anthropology, and Reflective Judgment* (Univ. Chicago Press, Chicago, 1999).
6. De Waal, F. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals* (Harvard Univ. Press, Cambridge, 1996).
7. Greene, J. D. & Haidt, J. How (and where) does moral judgment work? *Trends Cogn. Sci.* **6**, 517–523 (2002).
8. Casebeer, W. D. *Natural Ethical Facts: Evolution, Connectionism, and Moral Cognition* (MIT Press, Cambridge, Massachusetts, 2003).
9. Hume, D. *A Treatise of Human Nature* (Clarendon, Oxford, 1739/1985).
10. Moore, G. E. *Principa Ethica, Revised Edition* (Cambridge Univ. Press, New York, 1994).
11. Greene, J. From neural 'is' to moral 'ought': what are the moral implications of neuroscientific moral psychology? *Nature Rev. Neurosci.* **4**, 847–850 (2003).
12. Saver, J. L. & Damasio, A. R. Preserved access and processing of social kowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia* **29**, 1241–1249 (1991).
13. Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50**, 7–15 (1994).
14. Damasio, A. R. *Descartes' Error: Emotion, Reason, and the Human Brain* (G. P. Putnam and Sons, New York, 1994).
15. Anderson, S. W., Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neurosci.* **2**, 1032–1037 (1999).
16. Bauer, L. O. & Hesselbrock, V. M. P300 decrements in teenagers with conduct problems: implications for substance abuse risk and brain development. *Biol. Psychiatry* **46**, 263–272 (1999).
17. Moll, J. *et al.* The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* **22**, 2730–2736 (2002).
18. Moll, J., de Oliveira-Souza, R., Bramati, I. E. & Grafman, J. Functional networks in emotional moral and nonmoral social judgments. *Neuroimage* **16**, 696–703 (2002).
19. Hamann, S. B., Ely, T. D., Hoffman, J. M. & Kilts, C. D. Ecstasy and agony: activation of the human amygdala in positive and negative emotion. *Psychol. Sci.* **13**, 135–141 (2002).
20. Adolphs, R., Tranel, D. & Damasio, A. R. The human amygdala in social judgment. *Nature* **393**, 470–474 (1998).
21. Montague, P. R. *et al.* Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage* **16**, 1159–1164 (2002).
22. Van Veen, V. *et al.* Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage* **14**, 1302–1308 (2001).
23. Van Veen, V. & Carter, C. S. The timing of action-monitoring processes in the anterior cingulate cortex. *J. Cogn. Neurosci.* **14**, 593–602 (2002).
24. Bunge, S. A. *et al.* Prefrontal regions involved in keeping information in and out of mind. *Brain* **124**, 2074–2086 (2001).
25. MacDonald, A. W. III, Cohen, J. D., Stenger, V. A. & Carter, C. S. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* **288**, 1835–1838 (2000).
26. Baron-Cohen, S. *Mindblindness: An Essay on Autism and Theory of Mind* (MIT Press, Cambridge, Massachusetts, 1995).
27. Rizzolatti, G., Fogassi, L. & Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Rev. Neurosci.* **2**, 661–670 (2001).
28. Call, J. Chimpanzee social cognition. *Trends Cogn. Sci.* **5**, 388–393 (2001).
29. Adolphs, R. Cognitive neuroscience of human social behaviour. *Nature Rev. Neurosci.* **4**, 165–178 (2003).
30. Haidt, J. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* **108**, 814–834 (2001).
31. Greene, J. D. *et al.* An fMRI investigation of emotional engagement in moral judgment. *Science* **293**, 2105–2108 (2001).
32. Moll, J., de Oliveira-Souza, R. & Eslinger, P. J. Morals and the human brain: a working model. *Neuroreport* **14**, 299–305 (2003).
33. Fuster, J. M. *Cortex and Mind: Unifying Cognition* (Oxford Univ. Press, New York, 2002).