



How does morality work in the brain? A functional and structural perspective of moral behavior

Leo Pascual¹, Paulo Rodrigues^{1,2} and David Gallardo-Pujol^{1,3*}

¹ Department of Personality, University of Barcelona, Barcelona, Spain

² Mint Labs S.L., Barcelona, Spain

³ Institute for Brain, Cognition, and Behavior (IR3C), Universitat de Barcelona, Barcelona, Spain

Edited by:

Gordon M. Shepherd, Yale
University School of Medicine, USA

Reviewed by:

Antonio Pereira, Federal University
of Rio Grande do Norte, Brazil
Alexander J. Shackman, University
of Maryland, USA

*Correspondence:

David Gallardo-Pujol, Department of
Psychology, University of Barcelona,
Passeig de la Vall d'Hebron 171,
Barcelona, 08035, Spain
e-mail: david.gallardo@ub.edu

Neural underpinnings of morality are not yet well understood. Researchers in moral neuroscience have tried to find specific structures and processes that shed light on how morality works. Here, we review the main brain areas that have been associated with morality at both structural and functional levels and speculate about how it can be studied. Orbital and ventromedial prefrontal cortices are implicated in emotionally-driven moral decisions, while dorsolateral prefrontal cortex appears to moderate its response. These competing processes may be mediated by the anterior cingulate cortex. Parietal and temporal structures play important roles in the attribution of others' beliefs and intentions. The insular cortex is engaged during empathic processes. Other regions seem to play a more complementary role in morality. Morality is supported not by a single brain circuitry or structure, but by several circuits overlapping with other complex processes. The identification of the core features of morality and moral-related processes is needed. Neuroscience can provide meaningful insights in order to delineate the boundaries of morality in conjunction with moral psychology.

Keywords: fMRI, morality, neuroscience, moral judgement, social brain, neuroimaging

“By four-thirty in the morning the priest was all cleaned up. I felt a lot better. I always did, after. Killing makes me feel good. (...) It's a sweet release, a necessary letting go of all the little hydraulic valves inside. (...) It has to be done the right way, at the right time, with the right partner—very complicated, but very necessary.”

Dexter, Darkly dreaming Dexter (Jeff Lindsay)

Can immoral behavior sometimes turn out to be moral? What mechanisms underlie morality? The above quotation is taken from a scene in the American TV series “Dexter.” Dexter is a respected employee at the Miami Metro Police Department, and a family man. However, at night Dexter doubles as a covert serial killer, applying his own moral code and murdering assassins whom the legal system has failed to condemn or catch. To what extent can a murder be considered necessary or even moral? Dexter's code includes clear examples of moral paradoxes that are not yet well understood. Does Dexter's brain work in the same way as the brains of other people? Researchers in moral neuroscience have tried to find domain-specific structures and processes that shed light on what morality is and where it is in the brain, if in fact it is there at all.

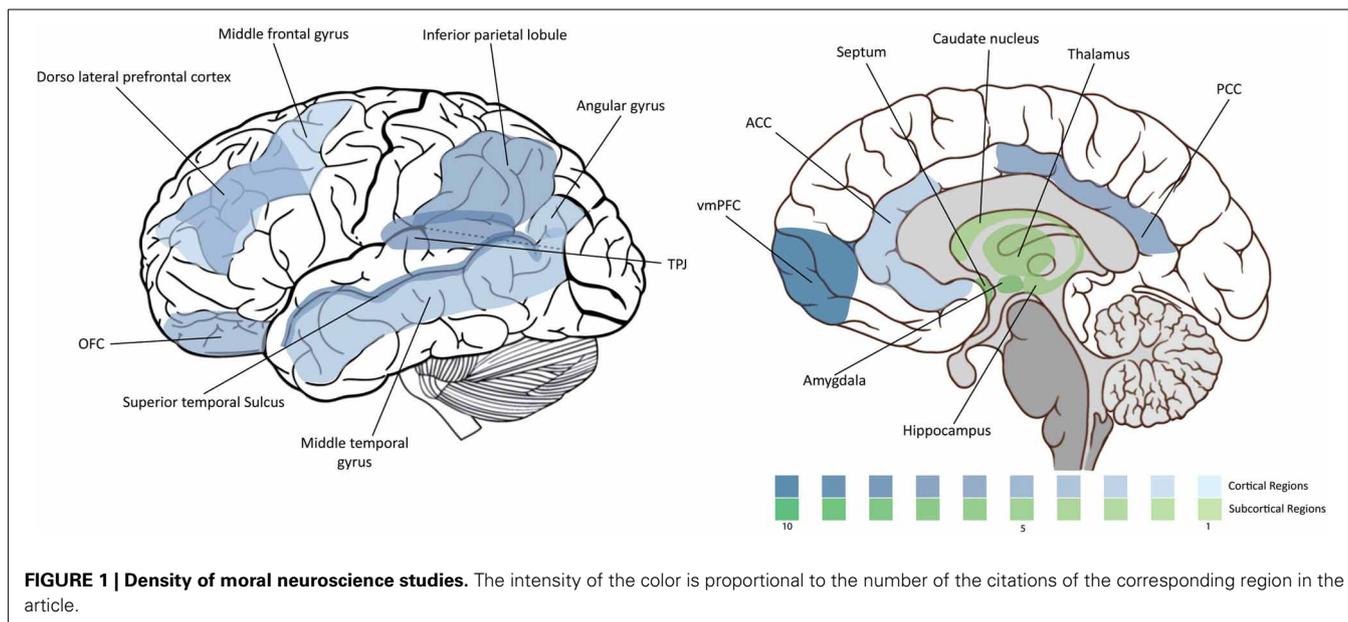
In this article, we focus on the history of the scientific study of neuroscience and the ways in which it has approached morality. We briefly review the main brain areas that have recently been associated with morality at both structural and functional levels and then discuss some of the implications of our review. We also speculate about how morality can be studied from the point of view of neuroscience. Here, we did a comprehensive review based on database search and references' search complemented with

Neurosynth as a tool to conduct reverse and forward inferences (Yarkoni et al., 2011).

WHAT IS MORALITY?

Morality has traditionally been regarded as a code of values guiding the choices and actions that determine the purpose and the course of our lives (Rand, 1964). Recently, it has been operationalized as a code of conduct that, given specified conditions, would be put forward by all rational persons (Gert, 2012). From a scientific point of view, the studies by Kohlberg represented a milestone in the psychological study of morality (Kohlberg, 1963, 1984). Kohlberg considered moral reasoning to be a result of cognitive processes that may exist even in the absence of any kind of emotions. However, findings in evolutionary psychology (Trivers, 1971; Pinker, 1997) and primatology (Flack and de Waal, 2000) suggested that emotions played a key part in the origins of human morality (e.g., kin altruism, reciprocal altruism, revenge).

Today, there is a general consensus in psychology and philosophy in favor of the differentiation of moral processes into two different classes: (1) rational, effortful and explicit, and (2) emotional, quick and intuitive (De Neys and Glumicic, 2008). The controversy remains, though, in how they interact. Among current models of moral processes and how they relate to each other, three distinct theories outstand (Greene and Haidt, 2002; Moll and Schulkin, 2009). The “*social intuitionist theory*” (Haidt, 2001) links research on automaticity (Bargh and Chartrand, 1999) to recent findings in neuroscience and evolutionary psychology. The “*cognitive control and conflict theory*” (Greene et al., 2004) postulates that responses



arising from emotion-related brain areas favor one outcome, while cognitive responses favor a different one (Kahneman and Frederick, 2007; McClure et al., 2007). According to the “*cognitive and emotional integration theory*,” behavioral choices cannot be split into cognitive vs. emotional. Complex contextual situations can make behavioral decisions exceptionally difficult (Gottfried, 1999; Moll et al., 2003).

HOW CAN MORALITY BE STUDIED SCIENTIFICALLY?

A variety of methods for exploring morality have been developed, from moral vs. non-moral situations to moral dilemmas (Young and Dungan, 2012). Moral dilemmas are situations in which every possible course of action breaches some otherwise binding moral principle (Thomson, 1985). The two main distinctions between moral dilemmas and judgments that have traditionally been taken into account are: (1) personal dilemmas and judgments, as opposed to impersonal ones (Greene et al., 2004); (2) utilitarian moral judgments vs. non-utilitarian ones (Brink, 1986).

These distinctions have led to the development of a variety of paradigms. Probably the most famous ones are the trolley paradigm (Thomson, 1985) and the footbridge dilemma (Navarrete et al., 2012). In both the trolley problem and the footbridge dilemma, the choice is between saving five people at the expense of killing one person or letting five die and one survive (Hauser, 2006; Greene, 2007). However, the latter meets the criteria of a personal dilemma, while the former does not (for extensive reviews of similar moral dilemmas, see Greene et al., 2004; Koenigs et al., 2007; Decety et al., 2011; Pujol et al., 2011). Other tasks that bring morality under experimental scrutiny present visual sentences or pictures (Greene et al., 2001; Harenski and Hamaan, 2006), or scales and questionnaires that can be used to assess moral behavior from a clinical point of view (see Rush et al., 2008 for a review).

THE NEUROANATOMY OF MORALITY

Being a highly complex process, morality involves a highly complex neural circuitry. In this section we overview the main brain areas and circuitry that have been associated with it. The “moral brain” comprises a large functional network that includes several brain structures. At the same time, many of these structures overlap with other regions that control different behavioral processes. We will review them in the following order: (1) the frontal lobe, (2) the parietal lobe, (3) the temporal lobe and insula, and (4) the subcortical structures. The findings are summarized in **Figure 1** and **Table 1**.

FRONTAL LOBE

The ventromedial prefrontal cortex (VMPFC) is consistently engaged in moral judgement (Greene et al., 2001; Moll et al., 2002; Harenski and Hamaan, 2006; Koenigs et al., 2007; Prehn et al., 2008; Harada et al., 2009). VMPFC seems to play a crucial role in the mediation of the emotions engaged during moral processing (Young and Koenigs, 2007). Patients with VMPFC lesions are reported to be significantly more likely to endorse utilitarian responses to hard personal moral dilemmas (Koenigs et al., 2007) and have trouble representing the abstract consequences of their decisions (Krajbich et al., 2009). It is also involved in adherence to social norms and values (Moll et al., 2005) and in the integration of representations of others’ intentions with their outcomes during social decision-making (Cooper et al., 2010). The left VMPFC shows higher activation in subjects with lower moral judgment competence when identifying norm violations (Prehn et al., 2008).

The orbitofrontal cortex (OFC) has been associated with morality, and has been implicated in the on-line representation of reward and punishment (O’Doherty et al., 2001; Shenhav and Greene, 2010). The right medial OFC was found to be activated during passive viewing of moral stimuli compared with non-moral stimuli (Harenski and Hamaan, 2006), while the activation

Table 1 | Studies that have addressed the morality problem from a neuroscientific viewpoint.

	VMPFC	OFC	DLPFC	ACC	MFG	IPR	TPJ	STS	A/MTG	AG	PCC	IC	HIP	AMG	THL	SPT	CN
Partiot et al. (1995)					X												
Bush et al. (2000)				X													
Allison et al. (2000)					X			X									
Greene et al. (2001)	X							X		X							
O'Doherty et al. (2001)		X						X									
Frith (2001)				X	X												
Moll et al. (2001)								X									
Farrow et al. (2001)									X								
Moll et al. (2002)	X	X						X			X						
Greene and Haidt (2002)						X					X						
Wicker et al. (2003)											X						
Greene et al. (2004)			X			X		X			X			X			
de Quervain et al. (2004)											X						X
Moll et al. (2005)	X										X						
Jackson et al. (2005)											X				X		
Harenski and Hamaan (2006)	X	X						X						X			
Prehn et al. (2008)																	
Borg et al. (2006)										X							
Völlm et al. (2006)											X						
Mendez (2006)											X						
Moll et al. (2006)																X	
Kent and Kiehl (2006)																X	
Luo et al. (2006)																	X
Koenigs et al. (2007)	X																
Young and Koenigs (2007)	X		X														
Ciarraido et al. (2007)							X										
Tsetsenis et al. (2007)													X				
Prehn et al. (2008)	X		X								X						
Haushofer and Fehr (2008)			X														
Harenski et al. (2008)						X		X									
Young and Saxe (2008)							X										
Buckholz et al. (2008)											X			X			
Hsu (2008)												X					
Harada et al. (2009)	X		X				X										
Krajcich et al. (2009)	X																
Greene and Paxton (2009)			X														
Fusar-Poli et al. (2009)										X			X				
Funk and Gazzaniga (2009)									X		X						

(Continued)

Table 1 | Continued

	VMPFC	OFC	DLPFC	ACC	MFG	IPR	TPJ	STS	A/MTG	AG	PCC	IC	HIP	AMG	THL	SPT	CN
Cooper et al. (2010)	X																
Shenhav and Greene (2010)		X															
Young et al. (2010)						X											
Cooper et al. (2010)							X										
Blair (2010)														X			
Sommer et al. (2010)						X											
Cáceda et al. (2011)							X										
Young and Dungan (2011)							X										
Moor et al. (2011)							X										
Sestieri et al. (2011)								X									
Decety et al. (2011)											X						
Immordino-Yang and Singh (2011)											X						
Total citations	10	4	6	3	2	5	6	7	3	3	6	10	3	6	3	2	2

VMPFC, ventromedial prefrontal cortex; OFC, orbitofrontal cortex; DLPFC, dorsolateral prefrontal cortex; ACC, anterior cingulate cortex; MFG, medial frontal gyrus; IPR, inferior parietal region; TPJ, temporo-parietal junction; STS, superior temporal sulcus; A/MTG, anterior/middle temporal gyrus; AG, angular gyrus; PCC, posterior cingulate cortex; IC, insular cortex; HIP, hippocampus; AMG, amygdala; THL, thalamus; SPT, septum; CN, caudate nucleus.

of the left OFC has been related to processing of emotionally salient statements with moral value (Moll et al., 2002). Greene et al. (2004) speculated that the orbital and ventromedial prefrontal cortices seem to be involved in emotionally driven moral decisions, whereas the dorsolateral prefrontal cortex (DLPFC) competes with it, eventually mitigating its responses (Greene et al., 2004). The DLPFC is differentially activated when subjects emit a utilitarian response (Young and Koenigs, 2007). This area is involved in cognitive control and problem-solving (Greene et al., 2004). The DLPFC plays an important role during the judgment of responsibility for crimes and its punishment from a third-party perspective (Haushofer and Fehr, 2008), and also in the analysis of situations that demand rule-based knowledge (Prehn et al., 2008). Greene and Paxton (2009) related it to lying processes, and others have hypothesized that it may trigger an executive function used to combine predictions based on social norms with inferences about the intent to deceive (Harada et al., 2009).

The anterior cingulate cortex (ACC) is involved in error detection (Shackman et al., 2011). It is activated when subjects generate a utilitarian response (Young and Koenigs, 2007). The ACC, among others, has been implicated in theory of mind (ToM) and self-referential tasks (Frith, 2001), and it has been involved in moral conflict monitoring (Greene et al., 2004, p. 391). The medial frontal gyrus is another frontal region that seems to intervene in ToM, and also in other social functions relevant to moral judgment (Amodio and Frith, 2006) and in the integration of emotion into decision-making and planning (Partiot et al., 1995).

PARIETAL LOBE

The inferior parietal region is mainly associated with working memory and cognitive control, and so its recruitment during moral processing might be due to some cognitive engagement (Greene et al., 2004; Harenski et al., 2008; Cáceda et al., 2011). One of its functions, together with the posterior area of the superior temporal sulcus (STS) which we will review below, seems to be the perception and representation of social information that may be crucial for making inferences about others' beliefs and intentions (Allison et al., 2000) and the representation of personhood (Greene and Haidt, 2002).

The temporo-parietal junction (TPJ) plays a key role in moral intuition and in belief attribution during moral processing in others (Young and Saxe, 2008; Harada et al., 2009; Young et al., 2010; Moor et al., 2011; Young and Dungan, 2011). The TPJ, as well as the precuneus, is involved in encoding beliefs and integrating them with other relevant features of the action such as the outcome (Young and Saxe, 2008). The right TPJ and the precuneus are active when subjects process prior intentions, while the left TPJ is activated when a subset of social intentions is involved (Ciaramidaro et al., 2007) as well as lying (Harada et al., 2009). The disruption of the right TPJ activity affects the capacity to use mental states during moral judgment (Young et al., 2010). In the dictator game, activation in the TPJ is associated with punishment of the excluders through lower offers (Moor et al., 2011).

TEMPORAL LOBE

Temporal lobe is one of the main neural regions activated during ToM tasks (Völlm et al., 2006; Ciaramidaro et al., 2007; Muller

et al., 2010). Structural abnormalities within this area have even been related to psychopathy (Blair, 2010; Pujol et al., 2011).

One of the main temporal sub-regions involved in moral judgment is the superior temporal sulcus (STS) (Allison et al., 2000; Moll et al., 2002; Greene et al., 2004; Harenski et al., 2008). This structure has been understood as an initial site of social perception (Allison et al., 2000) and has been repeatedly associated with emotional processing and social cognition (Greene et al., 2004; Harenski et al., 2008). The STS has been described as indispensable for making inferences about others' beliefs and intentions (Allison et al., 2000). Increased activity of this area is also observed in personal dilemmas compared with other types (Greene et al., 2001). In the dictator game the STS has been found to be activated when subjects applied punishment to the excluders (Moor et al., 2011). The posterior STS shows greater activity during justice-based dilemmas than in care-based dilemmas (Harenski et al., 2008). Subjects with lower moral judgment competence showed greater activation in the left posterior STS when identifying norm violations (Prehn et al., 2008).

The anterior/middle temporal gyrus has been also related to moral judgment (Moll et al., 2001; Greene et al., 2004; Harenski and Hamaan, 2006). Angular gyrus engagement has been observed during the evaluation of personal moral dilemmas (Greene et al., 2001; Borg et al., 2006; Funk and Gazzaniga, 2009).

LIMBIC LOBE

The posterior cingulate cortex (PCC) is known to be involved in the processing of personal memory, self-awareness and emotionally salient stimuli (Sestieri et al., 2011). It is one of the brain regions that exhibit greater engagement in personal than in impersonal dilemmas (Funk and Gazzaniga, 2009). Its activation has been related to social ability (Greene et al., 2004), empathy (Völlm et al., 2006) and forgiveness (Farrow et al., 2001), and can predict the magnitude of the punishments applied in criminal scenarios (Buckholz et al., 2008).

The insular cortex is also engaged in moral tasks (Moll et al., 2002; Greene et al., 2004). It exhibits greater activation in first-person and other-person experiences of disgust (Wicker et al., 2003). It is associated with emotional processing (Greene et al., 2004), empathic sadness in young subjects (Decety et al., 2011), detection and processing of uncertainty (Cooper et al., 2010) and perception of inequity (Hsu, 2008).

The anterior insular cortex is involved in visceral somatosensation, emotional feeling and regulation, and empathy (Immordino-Yang and Singh, 2011). This sub-region is activated during the experiencing of anger or indignation (Wicker et al., 2003; Moll et al., 2005), and when perceiving or assessing painful situations in others (Jackson et al., 2005). Its activation is also correlated with empathy scores (Völlm et al., 2006) and with unfair offers in a 'ultimatum game' (Hsu, 2008).

SUBCORTICAL STRUCTURES

The hippocampus is known to be a crucial region for the acquisition and retrieval of fear conditioning (Tsetsenis et al., 2007) and plays a facilitative role in inducing appropriate emotional reactions, in self-related processing during social

emotions (Immordino-Yang and Singh, 2011) and in the processing of emotional facial expressions (Fusar-Poli et al., 2009).

The amygdala is a necessary structure for moral learning (Mendez, 2006). It is involved in the evaluation of moral judgments (Greene et al., 2004) and in empathic sadness during morally-salient scenarios (Decety et al., 2011). It can predict punishment magnitude in criminal scenarios (Buckholz et al., 2008). Its dysfunction has been implicated in the affective deficits in psychopathy (Blair, 2010).

Rating empathic sadness, and perceiving and assessing painful situations has been associated with significant activation changes in the thalamus (Jackson et al., 2005; Decety et al., 2011). Bilateral thalamic activations are also observed when subjects are asked to choose between following a moral rule or a personal desire (Sommer et al., 2010).

The septum is activated while subjects make charitable contributions (Moll et al., 2006) and has been associated with psychopathy (Kent and Kiehl, 2006). Finally, the caudate nucleus is activated during altruistic punishment (de Quervain et al., 2004) and during the evaluation of morally salient stimuli (Luo et al., 2006).

DISCUSSION AND CONCLUSION

Moral neuroscience is an intricate and expanding field. This review summarizes the main scientific findings obtained to date. Morality is a set of complex emotional and cognitive processes that is reflected across many brain domains. Some of them are recurrently found to be indispensable in order to emit a moral judgment, but none of them is uniquely related to morality. The orbital and ventromedial prefrontal cortices are implicated in emotionally-driven moral decisions, whereas the dorsolateral prefrontal cortex seems to mitigate the salience of prepotent emotional responses. These competing processes may be monitored by the anterior cingulate cortex, which is also crucial for ToM. The TPJ and the STS play important roles in the attribution of others' beliefs and intentions. The insular cortex is engaged during empathic processes, and seems to be in charge of the evaluation of disgust and inequity. Other regions such as the posterior cingulate cortex, the anterior/middle temporal gyrus and the inferior parietal lobe seem to play a more complementary role in morality, being recruited in order to accomplish general cognitive processes engaged during the moral tasks proposed (e.g., working memory or cognitive control). On the other hand, regions like the amygdala seem to play an important role in the processing of emotions involved in moral judgment. Some of the emotions processed are more central to morality than others, but all emotions contribute to moral judgment given specific contextual situations.

The neural circuits of brain regions implicated in morality overlap with those that regulate other behavioral processes, suggesting that there is probably no undiscovered neural substrate that uniquely supports moral cognition. The most plausible option is that the "moral brain" does not exist *per se*: rather, moral processes require the engagement of specific structures of both the "emotional" and the "cognitive" brains, and the difference with respect to other cognitive and emotional processes may lie in the

content of these processes, rather than in specific circuits. Some authors, though, have related morality to basic emotions such as *disgust* (Chapman et al., 2009). Further research is needed in order to uncover the relationships between basic emotions and morality, as well as basic cognition blocks such as attentional control (van Dillen et al., 2012).

Given that morality is a highly complex process influenced by many factors, future studies should take into account individual differences (e.g., personality, genetics, religiosity, cultural and socioeconomic level) in order to understand the variety of mechanisms that govern it. Genetic factors and environmental-dependent processes during developmental stages may strengthen specific neural circuits that process various moral dimensions (Gallardo-Pujol et al., submitted).

Another important constraint in moral research is the heterogeneity of the tasks used in different studies to assess morality, which makes the comparison of the different results extremely difficult. Moreover, some of the tasks proposed barely suggest actual daily moral situations and usually require abstract evaluation, a circumstance that may blur the results obtained. The inclusion of innovative techniques such as

immersive virtual environments (Slater et al., 2006; Navarrete et al., 2012) adds apparent validity to moral dilemmas and may facilitate the generalization of results to real-life settings.

All in all, morality is supported not by a single brain circuitry or structure, but by a multiplicity of circuits that overlap with other general complex processes. One of the key issues that needs to be addressed is the identification of the core features of morality and moral-related processes. In this endeavor, neuroscience can provide meaningful insights in order to delineate the boundaries of morality in conjunction with moral psychology.

ACKNOWLEDGMENTS

This manuscript was made possible through the support of a grant from The Character Project at Wake Forest University and the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of The Character Project, Wake Forest University or the John Templeton Foundation. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

- Allison, T., Puce, A., and McCarty, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278. doi: 10.1016/S1364-6613(00)01501-1
- Amodio, D., and Frith, C. D. (2006). Meeting the minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–177. doi: 10.1038/nrn1884
- Bargh, J. A., and Chartrand, T. L. (1999). The unbearable automaticity of being. *Am. Psychol.* 54, 462–479. doi: 10.1037/0003-066X.54.7.462
- Blair, R. J. R. (2010). Psychopathy, frustration, and reactive aggression: the role of ventromedial prefrontal cortex. *Br. J. Psychol.* 110, 383–399. doi: 10.1348/000712609X418480
- Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., and Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: an fMRI investigation. *J. Cogn. Neurosci.* 18, 803–817. doi: 10.1162/jocn.2006.18.5.803
- Brink, D. O. (1986). Utilitarian morality and the personal point of view. *J. Philos.* 83, 417–438. doi: 10.2307/2026328
- Buckholtz, J. W., Asplund, C. L., Dux, P. E., Zald, D. H., Gore, J. C., Jones, O. D., et al. (2008). The neural correlates of third-party punishment. *Neuron* 60, 930–940. doi: 10.1016/j.neuron.2008.10.016
- Bush, G., Luu, P., and Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 6, 215–222. doi: 10.1016/S1364-6613(00)01483-2
- Cáceda, R., James, G. A., Ely, T. D., Snarey, J., and Kilts, C. D. (2011). Mode of effective connectivity within a putative neural network differentiates moral cognitions related to care and justice ethics. *PLoS ONE* 6:e14730. doi: 10.1371/journal.pone.0014730
- Chapman, H. A., Kim, D. A., Susskind, J. M., and Anderson, A. K. (2009). In bad taste: evidence for the oral origins of moral disgust. *Science* 323, 1222–1226. doi: 10.1126/science.1165565
- Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., et al. (2007). The intentional network: how the brain reads varieties of intentions. *Neuropsychologia* 45, 3105–3113. doi: 10.1016/j.neuropsychologia.2007.05.011
- Cooper, J. C., Kreps, T. A., Wiebe, T., Pirkel, T., and Knutson, B. (2010). When giving is good: ventromedial prefrontal cortex activation for others' intentions. *Neuron* 67, 511–521. doi: 10.1016/j.neuron.2010.06.030
- De Neys, W., and Glumicic, T. (2008). Conflict monitoring in dual process theories of reasoning. *Cognition* 106, 1248–1299. doi: 10.1016/j.cognition.2007.06.002
- de Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., and Buck, A. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.
- Decety, J., Michalska, K. J., and Kinzler, K. D. (2011). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cereb. Cortex* 22, 209–220. doi: 10.1093/cercor/bhr111
- Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarriner, N., et al. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport* 12, 2433–2438. doi: 10.1097/00001756-200108080-00029
- Flack, J., and de Waal, F. B. M. (2000). “Any animal whatever”: darwinian building blocks of morality in monkeys and apes. *J. Conscious. Stud.* 7, 1–29.
- Frith, U. (2001). Mind blindness and the brain in autism. *Neuron* 32, 969–979. doi: 10.1016/S0896-6273(01)00552-9
- Funk, C. M., and Gazzaniga, M. S. (2009). The functional brain architecture of human morality. *Curr. Opin. Neurobiol.* 19, 678–681. doi: 10.1016/j.conb.2009.09.011
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Sargu-ladze, S., et al. (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatr. Neurosci.* 34, 418–432.
- Gert, B. (2012). The definition of morality. *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University.
- Gottfried, K. (1999). Moral calculus and the bom. *Nature* 401, 117. doi: 10.1038/43571
- Greene, J., and Haidt, J. (2002). How (and where) does moral judgment work? *Trends Cogn. Sci.* 6, 517–523. doi: 10.1016/S1364-6613(02)02011-9
- Greene, J. D. (2007). “The secret joke of Kant’s soul,” in *Moral Psychology: The Neuroscience of Morality: Emotion, Disease, and Development*, Vol. 3, ed W. Sinnott-Armstrong (Cambridge, MA: MIT Press), 35–117.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., and Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400. doi: 10.1016/j.neuron.2004.09.027
- Greene, J. D., and Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12506–12511. doi: 10.1073/pnas.0900152106
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108. doi: 10.1126/science.1062872

- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295X.108.4.814
- Harada, T., Itakura, S., Xu, F., Lee, K., Nakashita, S., Saito, D. N., and Sadato, N. (2009). Neural correlates of the judgment of lying: a functional magnetic resonance imaging study. *Neurosci. Res.* 63, 24–34. doi: 10.1016/j.neures.2008.09.010
- Harenski, C. L., Antonenko, O., Shane, M. S., and Kiehl, K. (2008). Gender differences in neural mechanisms underlying moral sensitivity. *Soc. Cogn. Affect. Neurosci.* 3, 313–321. doi: 10.1093/scan/nsn026
- Harenski, C. L., and Hamaan, S. (2006). Neural correlates of regulating negative emotions related to moral violations. *Neuroimage* 30, 313–324. doi: 10.1016/j.neuroimage.2005.09.034
- Hauser, M. D. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York, NY: Ecco/HarperCollins Publishers.
- Haushofer, J., and Fehr, E. (2008). You shouldn't have: your brain on others' crimes. *Neuron* 60, 738–740. doi: 10.1016/j.neuron.2008.11.019
- Hsu, M. (2008). The right and the good: distributive justice and neural encoding of equity and efficiency. *Science* 320, 1092–1095. doi: 10.1126/science.1153651
- Immordino-Yang, M. H., and Singh, V. (2011). Hippocampal contributions to the processing of social emotions. *Hum. Brain Mapp.* 34, 945–955. doi: 10.1002/hbm.21485
- Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Image Process.* 24, 771–779.
- Kahneman, D., and Frederick, S. (2007). Frames and brains: elicitation and control of response tendencies. *Trends Cogn. Sci.* 11, 45–46. doi: 10.1016/j.tics.2006.11.007
- Kent, A., and Kiehl, K. (2006). A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. *Psychiatry Res.* 142, 107–128. doi: 10.1016/j.psychres.2005.09.013
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., and Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446, 908–911. doi: 10.1038/nature05631
- Kohlberg, L. (1963). *Moral Development and Identification*. *Child Psychology: The Sixty-Second Yearbook of the National Society for the Study of Education. Part 1*.
- Kohlberg, L. (1984). *Essays in Moral Development, Volume I: The Psychology of Moral Development*. New York, NY: Harper and Row.
- Krajčich, I., Adolphs, R., Tranel, D., Denburg, N. L., and Camerer, C. F. (2009). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J. Neurosci.* 29, 2188–2192. doi: 10.1523/JNEUROSCI.5086-08.2009
- Lindsay, J. (2004). *Darkly Dreaming Dexter*. New York, NY: Vintage Books.
- Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., and Blair, R. J. R. (2006). The neural basis of implicit moral attitude—an IAT study using event-related fMRI. *Neuroimage* 30, 1449–1457. doi: 10.1016/j.neuroimage.2005.11.005
- McClure, S. M., Botvinick, M. M., Yeung, N., Greene, J., and Cohen, J. D. (2007). “Conflict monitoring in cognition–emotion competition,” in *Handbook of Emotion Regulation*, ed J. J. Gross (New York, NY: Guilford), 204–226.
- Mendez, M. F. (2006). What frontotemporal dementia reveals about the neurobiological basis of morality. *Med. Hypotheses* 67, 411–418. doi: 10.1016/j.mehy.2006.01.048
- Moll, J., de Oliveira-Souza, R., and Eslinger, P. (2003). Morals and the human brain. *Neuroreport* 14, 299–305. doi: 10.1097/00001756-200303030-00001
- Moll, J., Eslinger, P. J., and de Oliveira-Souza, R. (2011). Frontopolar and anterior temporal cortex activation in a moral judgment task: preliminary functional MRI results in normal subjects. *Arq. Neuropsiquiatr.* 59, 657–664. doi: 10.1590/S0004-282X2001000500001
- Moll, J., Krueger, F., Zahn, R., Pardini, M., Oliveira-souza, R., and Graham, J. (2006). Human fronto–mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15623–15628. doi: 10.1073/pnas.0604475103
- Moll, J., Oliveira-souza, R. De, Eslinger, P. J., Bramati, I. E., Andreiuolo, P. A., and Pessoa, L. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral *Emotions* 22, 2730–2736.
- Moll, J., and Schulkin, J. (2009). Social attachment and aversion in human moral cognition. *Neurosci. Biobehav. Rev.* 33, 456–465. doi: 10.1016/j.neubiorev.2008.12.001
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., and Grafman, J. (2005). Opinion: the neural basis of human moral cognition. *Nat. Rev. Neurosci.* 6, 799–809. doi: 10.1038/nrn1768
- Moor, B. G., Güroğlu, B., Op de Macks, Z. A., Rombouts, S. A., Van der Molen, M. W., and Crone, E. (2011). Social exclusion and punishment of excluders: neural correlates and developmental trajectories. *Neuroimage* 59, 708–717. doi: 10.1016/j.neuroimage.2011.07.028
- Muller, F., Simion, A., Reviriego, E., Galera, C., Mazaux, J.-M., Barat, M., et al. (2010). Exploring theory of mind after severe traumatic brain injury. *Cortex* 46, 1088–1099. doi: 10.1016/j.cortex.2009.08.014
- Navarrete, C. D., McDonald, M. M., Mott, M. L., and Asher, B. (2012). Virtual morality: emotion and action in a simulated three-dimensional “Trolley Problem.” *Emotion* 12, 364–370.
- O’Doherty, J. O., Kringelbach, M. L., Rolls, E. T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102. doi: 10.1038/82959
- Partiot, A., Grafman, J., Sadato, N., Wachs, J., and Hallett, M. (1995). Brain activation during the generation of non-emotional and emotional plans. *Neuroreport* 10, 1269–1272.
- Pinker, S. (1997). *How the Mind Works*. New York, NY: Norton.
- Prehn, K., Wartenburger, I., Meriau, K., Scheibe, C., Goodenough, O. R., and Villringer, A. (2008). Individual differences in moral judgment competence influence neural correlates of socio-normative judgments. *Soc. Cogn. Affect. Neurosci.* 3, 33–46. doi: 10.1093/scan/nsm037
- Pujol, J., Batalla, I., Contreras-Rodríguez, O., Harrison, B. J., Pera, V., Hernández-Ribas, R., et al. (2011). Breakdown in the brain network subserving moral judgment in criminal psychopathy. *Soc. Cogn. Affect. Neurosci.* 7, 917–923. doi: 10.1093/scan/nsr075
- Rand, A. (1964). *The Virtue of Selfishness*. New York, NY: Penguin Books.
- Rush, J. A., First, M. B., and Blacker, D. (2008). *Handbook of Psychiatric Measures*. Washington, DC: American Psychiatric Publishing.
- Sestieri, C., Corbetta, M., and Romani, G. L., and Shulman, G. L. (2011). Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses. *J. Neurosci.* 31, 4407–4420. doi: 10.1523/JNEUROSCI.3335-10.2011
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., and Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167. doi: 10.1038/nrn2994
- Shenhav, A., and Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron* 67, 667–677. doi: 10.1016/j.neuron.2010.07.020
- Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., et al. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS ONE* 1:e39. doi: 10.1371/journal.pone.0000039
- Sommer, M., Rothmayr, C., Döhlhnel, K., Meinhardt, J., Schwerdtner, J., Sodian, B., et al. (2010). How should I decide? The neural correlates of everyday moral reasoning. *Neuropsychologia* 48, 2018–2026. doi: 10.1016/j.neuropsychologia.2010.03.023
- Thomson, J. J. (1985). The trolley problem. *Yale Law J.* 94, 1395–1415. doi: 10.2307/796133
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57. doi: 10.1086/406755
- Tsetsenis, T., Ma, X. H., Lo Iacono, L., Beck, S. G., and Gross, C. (2007). Suppression of conditioning to ambiguous cues by pharmacogenetic inhibition of the dentate gyrus. *Nat. Neurosci.* 10, 896–902. doi: 10.1038/nn1919
- van Dillen, L., van der Wal, R., and van den Bos, K. (2012). On the role of attention and emotion in morality: attentional control modulates unrelated disgust in moral. *Pers. Soc. Psychol. Bull.* 38, 1222–1231. doi: 10.1177/0146167212448485
- Völlm, B. A., Taylor, A. N. W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage* 29, 90–98. doi: 10.1016/j.neuroimage.2005.07.022
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula: the common neural basis of

- seeing and feeling disgust. *Neuron* 40, 655–664. doi: 10.1016/S0896-6273(03)00679-2
- Yarkoni, T., Poldrack, R., Nichols, T., Van Essen, D., and Wager, T. (2011). NeuroSynth: a new platform for large-scale automated synthesis of human functional neuroimaging data. *Front. Neuroinform. Conference Abstract: 4th INCF Congress of Neuroinformatics*. 8, 665–670. doi: 10.3389/conf.fninf.2011.08.00058
- Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., and Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6753–6758. doi: 10.1073/pnas.0914826107
- Young, L., and Dungan, J. (2011). Where in the brain is morality? Everywhere and maybe nowhere. *Soc. Neurosci.* 7, 1–10. doi: 10.1080/17470919.2011.569146
- Young, L., and Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Soc. Neurosci.* 7, 1–10. doi: 10.1080/17470919.2011.569146
- Young, L., and Koenigs, M. (2007). Investigating emotion in moral cognition: a review of evidence from functional neuroimaging and neuropsychology. *Br. Med. Bull.* 84, 69–79. doi: 10.1093/bmb/ldm031
- Young, L., and Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *NeuroImage* 40, 1912–1920. doi: 10.1016/j.neuroimage.2008.01.057
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 29 January 2013; accepted: 10 August 2013; published online: 12 September 2013.
- Citation: Pascual L, Rodrigues P and Gallardo-Pujol D (2013) How does morality work in the brain? A functional and structural perspective of moral behavior. *Front. Integr. Neurosci.* 7:65. doi: 10.3389/fnint.2013.00065
- This article was submitted to the journal *Frontiers in Integrative Neuroscience*. Copyright © 2013 Pascual, Rodrigues and Gallardo-Pujol. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.