Varieties of moral motivation: Empirical approaches

Mark Alfano, Macquarie University

Abstract

This chapter examines three recent empirical approaches to the study of moral motivation: moral foundations theory, deep pragmatism, and morality-as-cooperation. All three approaches conceptualize moral motivation as a suite of desires, emotions, sentiments, dispositions, values, and relationships that move people to think, judge, and act in accordance with morality. Moral foundations theory posits five or six basic foundations: care, fairness, loyalty, authority, sanctity, and sometimes liberty. People are thought to be emotionally attuned to each foundation, though some are more sensitive to certain foundations than to others. Deep pragmatism posits that people tend to be motivated not only to promote their own narrow self interest but also the interests of their group. This can be seen as a sort of proto-morality, which has the capacity to develop into full-blown morality as one begins to conceive of one's group in ever-broader terms. Finally, morality-as-cooperation is grounded in game theory and posits that cooperative solutions to non-zero-sum games are morally good, and that being motivated to seek out and implement such solutions is what it means to be morally motivated. Empirical evidence gathered in the last decade supports the morality-as-cooperation hypothesis, though more work is needed to further elaborate it and test its predictions.

Keywords: moral motivation, moral foundations, deep pragmatism, morality as cooperation

Word Count: 6062

Introduction

Moral motivation, as it is conceptualized in contemporary empirical psychology, consists in the mental states, dispositions, and relationships that move people to pursue, promote, and protect moral goods. Whereas philosophers have sometimes insisted that only one type of mental state (e.g., desire, or emotion) serves this function, psychologists tend to assume that moral motivation is multiply realizable. Psychological frameworks are specified functionally, such that whatever motivates moral conduct and cognition counts as moral motivation. For this reason, psychologists are usually indifferent as to which structures realize this function. Presumably, many different things do so in different contexts, from innate sentiments, to inherited social and cultural norms, to constructed institutions and technologies.[1]

Additionally, psychological approaches to the topic of moral motivation typically presume a multi-level social model, in which individuals are embedded in but can also come into conflict with groups, and in which groups are in turn embedded in but can also come into conflict with other groups or larger populations. In other words, there are two fronts of potential cooperation and conflict in which morality operates: between the self and the ingroup, and between the ingroup and outsiders and other groups. Morality helps us to navigate these two zones of potential cooperation and conflict. This functionalist understanding of moral motivation is less cognitively demanding and normatively thicker, in Williams's (1985) sense, than many popular philosophical approaches. Over the last half century, philosophers have tended to debate questions of moral motivation by talking about explicit, conscious judgments that something is right or wrong, obligatory or permissible, morally good or bad (e.g., Mackie 1977) – and perhaps even whether these judgments are connected with some overarching theory of the good life. While psychologists of course do not deny that people sometimes make such judgments, they are also interested in cases in which people have more inchoate attitudes that nevertheless motivate them to act and think in moral ways – regardless of whether they themselves would describe their own mentalizing and conduct in moral terms. In other words, psychologists are interested in *de re* moral motivation, whereas philosophers have tended to focus on *de dicto* moral motivation. For these reasons, many philosophical distinctions and debates, such as Humeanism/anti-Humeanism and internalism/externalism are uninteresting to psychologists.[2]

This chapter examines three recent empirical approaches to the study of moral motivation: moral foundations theory, deep pragmatism, and morality-as-cooperation. All three approaches conceptualize moral motivation as a suite of desires, emotions, sentiments, dispositions, values, and relationships that move people to think, judge, and act in accordance with morality. Moral foundations theory has posited five or six basic foundations: care, fairness,

---

[1] For additional background on this point, see Moody-Adams (1997), Wong (2006, 2014), Alfano (2016a, chapter 5), Flanagan (2016), and Schroeder et al. (2010).

[2] For more on these traditional philosophical distinctions and debates, see the previous chapters in this volume, as well as Rosati (2016). Also not covered in this chapter are the upshots for moral motivation from empirical research on sociopathy (see Roskies 2003, 2006) and autism (see Kennett 2002).

loyalty, authority, sanctity, and sometimes liberty. People are thought to be emotionally attuned to each foundation, though some are more sensitive to certain foundations than to others. Deep pragmatism posits that people tend to be motivated not only to promote their own narrow self interest but also the interests of their group. This can be seen as a sort of proto-morality, which has the capacity to develop into full-blown morality as one begins to conceive of one's group in ever-broader terms. Finally, morality-as-cooperation is grounded in game theory and posits that cooperative solutions to non-zero-sum games are morally good, and that being motivated to seek out and implement such solutions is what it means to be morally motivated. Both philosophical argumentation and empirical evidence gathered in the last decade support the morality-as-cooperation hypothesis over its rivals, though more work is needed to further elaborate it and test its predictions.

Moral foundations theory

Moral foundations theory grew out of previous anthropological work by Shweder et al. (1997), according to which moral judgments and motivations across cultures tend to be elicited by a range of emotions.[3] Shweder and colleagues initially postulated that three emotions in particular could account for much of the variation in moral attitudes within and between cultures: contempt, anger, and disgust. Contempt governs hierarchical relationships. People whose conduct is (or is seen to be) unworthy or above their station are subject to the contempt of their community and the shame that this induces (Mason 2018). Anger is a more horizontal emotion: it is directed at those who have harmed oneself or those one is close to. It calls for punishment, or, in happier cases, a sense of guilt in the offender which results in apology and forgiveness (Strawson 2008). Disgust protects purity, arguably as a manifestation of the behavioral immune system. It signals that something or someone is impure and to-be-avoided. In basic cases, the target of disgust is infectious or poisonous, but the disposition to feel disgust can also be enculturated to target violations of moral norms, foreigners, and just about anything else (Strohminger & Kumar 2018, Kelly 2011).

In subsequent work, Haidt and colleagues expanded on the contempt-anger-disgust framework to posit five (later six) "foundations" for moral judgment and motivation: care, fairness, loyalty, authority, sanctity, and eventually liberty (Graham et al. 2011). Each of these foundations is conceptualized as a sentiment – a standing disposition to token relevant emotions when norms associated with the foundation are violated or notably adhered to. For instance, when someone experiences or witnesses a serious harm (violation of the care foundation), they are likely to respond with anger at the perpetrator; conversely, when someone experiences or witnesses a supererogatory act of care (noteworthy action in accordance with the care

[3] Note that this study was also conducted in the bad old days before the replication crisis forced psychological researchers to reform their methods, and thus had a total N of just 47 participants. In recent years, there has been a trend towards having hundreds, thousands, and in some cases tens of thousands of participants.

foundation), they are likely to respond with appreciation, praise, and perhaps even the emotion of elevation. The care foundation is a descendant of Shweder's anger-related disposition; the loyalty and authority foundations are descendants of Shweder's contempt-related disposition; and the sanctity foundation is derived from his thinking about disgust.

Clearly there is something to the connection between moral motivation and emotions, as philosophers from David Hume to Adam Smith to Friedrich Nietzsche to Franz Brentano have emphasized for centuries. However, it is not clear that moral foundations theory, which is notoriously unsystematic, draws all and only the right connections between emotions and morality. Despite being the most well-known of the three frameworks discussed in this chapter (Haidt has been cited nearly 90,000 times as of the writing of this chapter), moral foundations theory has many significant flaws and shortcomings. For instance, fairness in moral foundations theory is not conceptualized coherently. The foundation seems to include equity, equality, procedural fairness, which can of course come apart from one another. Second, moral foundations theory conflates fairness with reciprocity, which is a distinct phenomenon and modeled differently in game theory (see the section on morality-as-cooperation for more). Third, despite its name, moral foundations theory is less a theory and more a grab-bag collection. Haidt has admitted that the sanctity foundation is an "odd" fit (Haidt & Joseph 2004). It wasn't until 2013 that Graham et al. (2013) even attempted to specify criteria that qualify something as a specifically *moral* foundation, and they have never used these criteria to add or eliminate a single foundation, making the criteria seem more like post hoc confabulations. Indeed, Haidt has gone so far as to endorse what he calls an "ad hoc" approach (Haidt & Joseph 2011). Fourth, moral foundations theory is clearly not comprehensive because it lacks any place for kin altruism and moral obligations to family. Fifth, the foundations are supposed to be "modular" and independent of one another (Haidt & Joseph 2007), but subsequent cross-cultural research has shown that they cannot typically be teased apart in a large cross-cultural sample, indicating that the originally-posited five dimensions are too many (Iurino & Saucier 2020).

Perhaps the most egregious aspect of moral foundations theory is the inclusion – starting in a non-peer-reviewed popular monograph (Haidt 2012) – of the liberty foundation.[4] According to Haidt, adding this dimension was warranted to help distinguish traditional conservatives from libertarians in the United States. The empirical basis for this distinction is poor, and it does not replicate outside of the United States (Nejat et al. 2015; Yalçındağ et al. 2019). Given Haidt's emphasis on the global context in his earlier work, it is odd that this flimsy parochial construct would be seen as meriting its own foundation in an allegedly universal theory of human morality. The supposed evolutionary origin of this foundation in hypervigilance against domination and coercion is not well established and arguably conflicts with the allegedly distinct foundations of authority and loyalty. Furthermore, people who score high on this dimension often exhibit counter-moral attitudes and behaviors, such as resisting vaccination (Amin et al. 2017). In

[4] Iyer et al. (2012) subsequently published a peer-reviewed paper aimed at justifying this speculative new foundation. It includes an inordinate amount of praise for the moral imbecile Ayn Rand and her devotees.

addition, they often seem perfectly happy to support coercive policies as long as those policies are consistent with their narrow self-interest or their group identity. One could be forgiven for speculating that Haidt introduced the liberty foundation and began positioning himself publicly as a centrist critical of the "coddling of the American mind" (2018) in order to attract attention and funding from right-wingers in the United States after he moved from an academic position at the University of Virginia to the NYU Stern School of Business in 2012 (the year *The Righteous Mind* was published).

Deep pragmatism

In *Moral Tribes: Emotion, Reason, and the Gap between Us and Them,* a landmark book summarizing his own and others' research in moral psychology, Joshua Greene (2013) argues for a version of utilitarianism that he calls "deep pragmatism." Greene's book begins with the distinction mentioned above between two domains of possible cooperation and conflict: between the self and others, and between one's group and other groups. Greene argues that humans have both biologically and culturally evolved dispositions, technologies, and institutions that systematically help us to navigate the former domain of cooperation and conflict. For instance, most humans are innately horrified by the prospect of killing conspecifics; our hypersociality would be impossible without this aversion and arguably derives from a process of accidental self-domestication in our evolutionary history (Hare et al. 2012; Hare 2017). Indeed, this is such a strong, emotionally-inflected sentiment that much of modern military training aims to eradicate it – sometimes leading morally desensitized troops to commit atrocities (Trivigno 2013). This aversion to violence is, in the vast majority of humans, further enhanced during development as children are rewarded for being gentle towards and punished for hurting others – at least towards members of their families and ingroups (Nichols et al. 2016). Thus, already, we see that most people have at least some moral motivation.

Greene goes on to document empirical work that establishes that many emotionally-tinged moral motivations, such as gratitude, friendship, and vengefulness are widespread if not completely universal among humans. Gratitude and friendship are examples of positive reciprocity, whereas vengefulness is an example of negative reciprocity. Both forms of reciprocity are evolutionarily stable cooperative strategies under conditions that are frequently encountered by humans, meaning that societies that settle on equilibria involving these strategies are likely to continue with them unless they encounter some exogenous disruption. Greene also discusses loyalty, which can be part of a dyadic friendship, but which can also connect people to larger groups. Loyalty enables collective action, which empowers groups to accomplish much more than any of their individual members could hope to do on their own (think of Rousseau's stag hunt). And larger, more trusting groups are able to accomplish more than smaller, less trusting groups, creating evolutionary pressure to ramp up loyalty even more (Dunbar 1991, 1992, 1993, 2005; see also Alfano 2016a). This pressure is reflected in moral norms that encourage dilation of the circle of moral concern; such norms are explicitly endorsed in historical

philosophy in both in Western philosophy (e.g., Cicero's concentric circles of moral concern) and in Eastern philosophy (e.g., Mengzi on the virtue of *ren* or compassion), and have been endorsed by contemporary utilitarians such as Singer (2011), who uses the metaphor of an expanding circle of moral concern that includes not just humans but all sentient life.

Loyalty is especially salient in the context of intergroup conflict – the second domain in which cooperation and conflict is possible. Larger groups that trust in their members' loyalty are more capable of both resisting and dominating smaller groups and groups whose members don't trust each others' loyalty. This may seem to put limits on the extent to which the circle of moral concern can reasonably be expanded. Is the utopia of universal benevolence, in which every human or even every sentient being is part of the ingroup, a feasible prospect? Greene worries that the very same emotional dispositions that encourage us to cooperate with our ingroups may exacerbate intergroup conflict and violence – a problem that he dubs the tragedy of commonsense morality. This problem sometimes manifests in what other researchers have termed "virtuous violence," where people engage in aggressive, even genocidal, behavior that they themselves perceive as morally good because it supports their ingroup at the expense of others (Fiske & Rai 2014).

Importantly for Greene, these dispositions are thought to issue in swift, automatic, hard-to-control, and non-introspectable affects and emotions (what's sometimes called System 1 processes). He likens such affects and emotions to processes studied in cognitive science using paradigms such as the Stroop task. In a typical Stroop task, participants are asked to read words printed in different fonts. They tend to be slower and less accurate when the meaning of the word (e.g., 'red') differs from the font in which it's printed (e.g., blue font): perceptual and semantic processes (which are also System 1) interfere with each other. According to Greene et al. (2001), a similar sort of interference occurs when fast, emotional processes that typically result in apt moral judgments misfire, leading people to make judgments that they would reject on rational reflection. He calls this a moral Stroop effect, and uses seemingly inconsistent responses to the trolley problem (Foot 1967) as evidence. In particular, Greene contends that people reject sacrificing one to save many in the footbridge version of the trolley problem because their automatic emotional reactions get the better of them, whereas they engage in cool, unemotional reasoning in the switch case to arrive at the morally correct judgment that sacrificing one to save many is acceptable. Unfortunately for Greene, the analogy between the original Stroop task and his alleged moral Stroop task does not bear critical scrutiny. In particular, the timescale is untenable (Huebner 2015). Automatic, System 1 emotional responses in the amygdala and ventromedial prefrontal cortex occur within 200 milliseconds (Decety & Cacioppo 2012). Reading the footbridge version of the trolley dilemma takes about 20 seconds – two orders of magnitude more. There is more than enough time to deliberate while reading these vignettes, meaning that participants' responses are not explicable by appeal to the System 1 / System 2 distinction.

From a more theoretical point of view, it's also unclear why we should draw a hard line between arational, automatic, emotional processes, on the one hand, and deliberative, rational

processes on the other hand. It's more plausible that emotions are one way in which humans implement their practical rationality, both in the moral domain and elsewhere (Railton 2014; Alfano 2016b, chapter 5). Beyond this, Greene's suggestion that the tragedy of commonsense morality is about abstract ideological disagreements between groups about how best to organize society, rather than attempts at and resistance to domination between groups, is empirically and politically untenable. Most people do not have a fleshed-out ideology and are perfectly happy to modify whatever piecemeal ideologies they've adopted to keep in lockstep with their ingroup.[5]

Finally, Greene's so-called deep pragmatism ends up sacrificing its own empirical foundation when he shifts from an endorsement of cooperation – which only occurs in the context of nonzero-sum interactions (discussed in more detail in the next section) to a version of hedonic utilitarianism. Universal benevolence, sadly, is not an evolutionarily stable strategy, which means that deep pragmatism if it were ever implemented at a global scale would be unlikely to stick around for long – at least, as long as resources remain sufficiently scarce, which seems likely in the face of impending climate change.. As Bicchieri (2016) has shown, it's often better to implement a stable, suboptimal strategy than to shoot for the stars and fall into the gutter.

Morality as cooperation

As mentioned above, most psychological research presupposes or suggests that the function of morality is to promote cooperation (Curry 2016; Rai & Fiske 2011; Sterelny & Fraser 2016; Tomasello & Vaish 2013).[6] According to the theory of morality-as-cooperation, humans face, and have faced, a range of different 'nonzero-sum' problems of cooperation, and have evolved and invented a range of solutions to them. A game is a strategic situation in which the value of an outcome for an individual depends not only on what they choose to do but also on what others choose to do. In a zero-sum game, one player wins at the other's expense, but in *nonzero*-sum games, it is possible for one individual to be better off without that benefit coming at the expense of someone else (it can be a win-win). A classic example of a nonzero-sum interaction is Rousseau's stag hunt, as mentioned above. Cooperative solutions to nonzero-sum games take a variety of forms, including character traits, strategies, dispositions, behaviors, rules, norms, institutions, and technologies. Together, they motivate cooperative behavior, and provide the criteria by which we judge the behavior, attitudes, and traits of ourselves and others (Curry 2016).

What's more, because there are many types of cooperation, morality-as-cooperation predicts that there will be many dimensions or aspects of morality. For example, drawing on evolutionary game theory, Curry (2016) has argued that there are at least seven distinct types of cooperation: (1) the allocation of resources to kin; (2) coordination to mutual advantage; (3)

---

[5] See, for example, American Republicans' flip-flopping on the importance of moral character in political leaders in the wake of Donald Trump's rise (url = < https://news.gallup.com/poll/235022/presidential-moral-leadership-less-important-republicans.aspx, accessed 28 February 2022).

[6] Note that, on this very abstract point both Greene (2015, p. 40) and Haidt & Kesebir (2010, p. 800) agree. The devil is of course in the details.

social exchange; and conflict resolution through contests featuring (4) hawkish displays of dominance and (5) dovish displays of deference; (6) division of disputed resources; and (7) recognition of prior possession. And each of these types of cooperation gives rise to a corresponding type of morality: (1) family values, (2) group loyalty, (3) reciprocity, (4) heroism, (5) deference, (6) fairness, and (7) property rights.

> 1. Family Values: Kin selection explains why we love and care for our families, why parents feel a duty of care to their children, why we feel a special obligation to help our extended families, and why we abhor incest (Hamilton 1964; Lieberman, Tooby, & Cosmides 2003).
>
> 2. Group Loyalty: Mutualism explains why we coordinate our activities to pursue projects of mutual interest (e.g., the stag hunt), why we form groups, clubs, and coalitions, why we value these groups, their members, and our membership in them, why we adopt local norms and conventions, why we feel an obligation to come to the aid of group members, and why we value loyalty, unity, and solidarity (Lewis 1969; Royce 1908).
>
> 3. Reciprocity: Social exchange explains why we seek opportunities for mutually-beneficial trade, and why we feel we ought to trust others, return favors (positive reciprocity), keep promises, pay debts, fulfill contracts, be grateful for favors received, feel guilty for favors not returned, avenge injuries (negative reciprocity), punish cheats, apologize for causing injuries, and forgive those who apologize (Cosmides & Tooby 2005; Trivers 1971).
>
> 4. & 5. Heroism and Deference: Conflict resolution explains why we minimize the mutual costs of disputes by engaging in ritual contests: why we proudly display cues of power and high status (excellences such as bravery and generosity); and, when bested, why we express humility, and respect, defer, and submit to our superiors (Curry 2007; Sznycer et al. 2017).
>
> 6. Fairness: Conflict resolution also explains why we resolve disputes over divisible resources by dividing or sharing them (often resulting in equal shares), and hence why we feel an obligation to negotiate, seek compromise, and be fair (Messick 1993; Skyrms 1996).
>
> 7. Property Rights: Conflict resolution also explains why we resolve disputes by recognising prior possession, hence why we feel we ought to respect others' property and territory, and refrain from theft (Gintis 2007).

According to morality-as-cooperation, each of these types of cooperation is valuable and valued, and it is these values that we call moral. Consequently, acting morally consists in acting in accordance with these cooperative strategies in order to realize mutually-beneficial cooperative opportunities. And moral motivation consists in being moved to act in these ways, regardless of whether one explicitly conceptualizes one's conduct as moral; indeed, one might even consciously think of one's actions as immoral while doing the right thing, as Nomy Arpaly

(2002) has explored in her discussion of the moral psychology of the character of Huck Finn as he is helping Jim, a runaway slave.

Morality as cooperation makes a number of predictions that have been supported by previous research. For example, it predicts that: there will be as many different types of morality as there are types of cooperation (and their combinations) (Curry et al. 2021); that each of these distinct types of morality will be the product of genetically-distinct, domain-specific psychological mechanisms (Zakharin et al. 2021); that each of these types of cooperation will be considered morally relevant, and morally good, and will constitute a distinct facet or factor (Curry et al. 2019a); and that each of these types of cooperation will always be considered morally good, in all cultures. To test the final prediction, Curry et al. (2019b) analyzed 600 ethnographic accounts of ethics from 60 societies, as recorded in the Human Relations Area Files. They found, first, that each of the seven cooperative behaviors were considered morally good in 99.9% of cases. Second, they found examples of most of these morals in most societies – crucially, there were no counter-examples, no societies in which any of these behaviors were considered morally bad. And third, they observed these morals with equal frequency across continents; they were not the exclusive preserve of 'the West' or any other region.

In light of the sophisticated mathematical foundation of morality-as-cooperation in evolutionary game theory and its recent demonstrated predictive successes across the globe, the morality-as-cooperation hypothesis is a clear frontrunner in the empirical quest to characterize moral motivation. However, this theory is not without its own challenges. First, because it is grounded in game theory, which models people as constantly making choices and acting, it presupposes a highly agentic conception of morality. What, we might ask, about moral patients? Are they beyond the remit of morality if their agency is partially or fully impaired?[7] This is not an idle question, as we typically think that we have moral obligations to (among others) people who suffer from severe mental and physical impairment, as well as orphaned babies who are not in a position to benefit from kin altruism. These humans are arguably incapable of acting cooperatively. Relatedly, many would say that we have moral obligations to non-human animals. When they are domesticated animals – especially dogs that have coevolved with humans – perhaps it is possible to make the argument that they are cooperative partners and thus deserve moral consideration.[8] But in the case of wild animals this argument seems implausible. If one is moved by Bentham's insistence that "the question is not, Can they *reason?* nor, Can they *talk?* but, Can they *suffer?*" (1789/1961, ch. 17, fn.) then one needs a response to this worry. The proponent of morality-as-cooperation could respond by saying that our moral obligations to the infirm, orphans, and non-human animals – and the urges most people clearly feel to respect these obligations – are derivative of other, cooperation-based values and dispositions that we acquire in the course of moral development. Children who torture domesticated or wild animals that they've captured are liable to go on to harm humans. Those who would abandon orphans and severely impaired elders may do the same to those with less severe debilities. So, to ensure that we do not fall down such slippery slopes, we have moral norms that, strictly speaking, grant moral status to humans and non-human animals that are incapable of cooperation. This response, however, may seem unsatisfying, insensitive, and instrumentalizing.

---

[7]As noted above, both moral foundations theory and deep pragmatism are also cooperative theories. They make less systematic use of game theory, however, and thus might find other ways to dodge this criticism.

[8] Do they also have moral obligations? That would be a strange result.

A second challenge faced by the theory of morality-as-cooperation relates to the problem of universal benevolence discussed in the previous section. We can approach by asking about zero-sum interactions: are all moral norms suspended in such situations? These situations may be rare in contemporary societies, but surely they exist sometimes. Arguably, they constitute a Hobbesian states of nature, and morality really does have no normative force in the state of nature (Curry et al. 2019b). However, the proponent of morality-as-cooperation could also respond (in a way resembling Hobbes himself) that if someone finds themselves in a zero-sum interaction, then they should have a higher-order motivation: not to cooperate (since cooperation is by definition impossible in a zero-sum interaction) but to shift the state of play such that cooperation becomes possible. To my knowledge no empirical work has been conducted to examine how prevalent such a second-order motivation. Such a second-order moral motivation could also help to shift mixed-motive nonzero-sum games (e.g., prisoner's dilemma, hawk/dove) to fully cooperative ones (e.g., stag hunt), thereby making cooperation both more likely and more efficient. Perhaps a bit grandiosely, we can call this second-order moral motivation a drive for moral progress, a possibility discussed at length in the context of evolutionary game theory by Kumar & Campbell (2022).

Conclusion

The three main empirically-informed theories of moral motivation are sanguine – perhaps more sanguine than many philosophers. All three theories agree that moral motivation is prevalent in humans, often at an early age. They disagree about exactly which structures embody such motivation and how they operate in our daily lives. Moral foundations theory posits sentiments that connect to domains such as care and purity. Deep pragmatism posits a range of automatic, affective processes that lead us to make moral judgments and decisions. Morality as cooperation posits a wide range of motivational structures that are united in moving us to cooperate in nonzero-sum interactions. These three theories also disagree more theoretically about what underpins the posited motivations. Moral foundations theory is proudly ad hoc. Deep pragmatism is an empirically-informed version of hedonistic utilitarianism. Morality as cooperation is rooted in game theory. Each view has its merits, but I hope to have shown in this chapter that morality as cooperation is the most promising of the three.

References

Alfano, M. (2016a). The topology of communities of trust. *Russian Sociological Review*, 15(4): 30-56.

Alfano, M. (2016b). *Moral Psychology: An Introduction*. Polity.

Amin, A. B., Bednarczyk, R. A., Ray, C. E., Melchiori, K. J., Graham, J., Huntsinger, J. R., & Omer, S. B. (2017). Association of moral values with vaccine hesitancy. *Nature Human Behaviour*, *1*(12), 873-880.

Arpaly, N. (2002). *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford University Press.

Bentham, J. (1789/1961). *An Introduction to the Principles of Morals and Legislation*. Garden City, NY: Doubleday.

Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.

Cosmides, L., & Tooby, J. (2005). Social exchange: The evolutionary design of a neurocognitive system. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (Vol. III, pp. 1295-1308). Cambridge, MA: MIT Press.

Curry, O. S. (2007). The conflict-resolution theory of virtue. In W. P. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. I, pp. 251-261). Cambridge, Massachusetts: MIT Press.

Curry, O. S. (2016). Morality as Cooperation: A Problem-Centred Approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The Evolution of Morality* (pp. 27-51): Springer International Publishing.

Curry, O., Alfano, M., & Brandt, M. (2021). Moral molecules: Morality as a combinatorial system. *Review of Philosophy and Psychology*.

Curry, O., Chesters, M., & Van Lissa, C. (2019a). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *Journal of Research in Personality*, 78: 106-24.

Curry, O., Mullins, D., & Whitehouse, H. (2019b). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1): 47-69.

Decety, J. & Cacioppo, S. (2012). The speed of morality. *Journal of Neurophysiology*, 108(11): 3068–72.

Dunbar, R. (2005). Gossip in evolutionary perspective. *Review of General Psychology*, 8: 100-110.

Dunbar, R. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16(4): 681-735.

Dunbar, R. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6): 469-93.

Dunbar, P. (1991). Functional significance of social grooming in primates. *Folia Primatologica*, 57: 121-31.

Fiske, A. P., & Rai, T. S. (2014). *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge University Press.

Flanagan, O. (2016). *The Geography of Morals: Varieties of Moral Possibility*. Oxford University Press.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5: 5–15.

Gintis, H. (2007). The Evolution of Private Property. *Journal of Economic Behavior & Organization*, 64(1), 1-16. Retrieved from http://dx.doi.org/10.1016/j.jebo.2006.02.002

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology* (Vol. 47, pp. 55-130). Academic Press.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366–385.

Greene, J. D. (2015). The rise of moral cognition. *Cognition*, 135(0), 39-42. doi:http://dx.doi.org/10.1016/j.cognition.2014.11.018

Greene, J., Sommerville, R., Nystrom, L., Darley, J., & Cohen, J. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293: 2105–08.

Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Pantheon.

Haidt, J., & Lukianoff, G. (2018). *The Coddling of the American Mind: How Good Intentions and Bad Ideas are Setting up a Generation for Failure*. Penguin UK.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55-66.

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The innate mind*, *3*, 367-391.

Haidt, J., & Joseph, C. (2011). How moral foundations theory succeeded in building on sand: A response to Suhler and Churchland. *Journal of Cognitive Neuroscience*, *23*(9), 2117-2122.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, G. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., pp. 797-832). Hoboken, NJ: Wiley.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7, 1-16, 17-52. doi:10.1016/0022-5193(64)90039-6

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annual review of psychology*, *68*, 155-186.

Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573-585.

Huebner, B. (2015). Do emotions play a constitutive role in moral cognition?. *Topoi*, *34*(2), 427-440.

Human Relations Area Files. (n.d.). eHRAF User Guide. url = <https://ehrafworldcultures.yale.edu/webhelp/#Get_Started/getting_started.html>, accessed 11 Feb 2021.

Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the Moral Foundations Questionnaire across 27 countries. *Assessment*, 27(2), 365-372.

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS ONE*, 7(8): e42366.

Kelly, D. (2011). *Yuck!: The Nature and Moral Significance of Disgust*. MIT press.

Kennett, J. (2002). Autism, empathy and moral agency. *Philosophical Quarterly*, 52 (208): 340-357.

Lewis, D. K. (1969). *Convention: a philosophical study.* Cambridge, Massachusetts: Harvard University Press.

Kumar, V. & Campbell, R. (2022). *A Better Ape: The Evolution of the Moral Mind and How it Made us Human*. Oxford University Press.

Lieberman, D., Tooby, J., & Cosmides, L. (2003). Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 270(1517), 819-826. doi:Doi 10.1098/Rspb.2002.2290

Mackie, J. (1977). *Ethics: Inventing Right and Wrong*. Penguin.

Mason, M. E. (2018). *The moral psychology of contempt*. Rowman & Littlefield.

Messick, D. M. (1993). Equality as a decision rule. In B. Mellers & J. Baron (Eds.), *Psychological Perspectives on Justice* (pp. 11-31). Cambridge: Cambridge University Press.

Moody-Adams, M. (1997). *Fieldwork in Familiar Places: Morality, Culture, and Philosophy*. Harvard University Press.

Nejat, P., Bagherian, F., Shokri, O., & Hatami, J. (2015, April). Mental representation of morality: The endorsement pattern of moral foundations among Iranians. In *2015 Sixth International Conference of Cognitive Science (ICCS)* (pp. 72-76). IEEE.

Nichols, S., Kumar, S., Lopez, T., Ayars, A., & Chan, H. Y. (2016). Rational learners and moral rules. *Mind & Language*, *31*(5), 530-554.

Otterbein, C. S. (1978). Methods in cross-cultural psychology. *American Psychologist, 33*(9), 857.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, 118(1), 57-75. doi:10.1037/a0021867

Railton, P. (2014). The affective dog and its rational tale: Intuition and attunement. *Ethics*, *124*(4), 813-859.

Rosati, Connie S. (2016). Moral motivation, *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/moral-motivation/>.

Roskies, A. (2003). Are ethical judgments intrinsically motivational? Lessons from 'acquired sociopathy', *Philosophical Psychology*, 16: 51–66.

Roskies, A. (2006). Patients with ventromedial frontal damage have moral beliefs, *Philosophical Psychology*, 19: 617–627.

Royce, J. (1908). *The Philosophy of Loyalty.* New York: Macmillan.

Schroeder, T., Roskies, A., and Nichols, S. (2010). Moral motivation, in J. Doris (ed.), *The Moral Psychology Handbook,* 72–110. Oxford University Press,.

Shweder, R., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. *Morality and health*, 119.

Singer, P. (2011). *The Expanding Circle: Ethics, Evolution, and Moral Progress*. Princeton University Press.

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Sterelny, K., & Fraser, B. (2016). Evolution and Moral Realism. *British Journal for the Philosophy of Science,* 68(4), 981-1006. Retrieved from https://academic.oup.com/bjps/article-abstract/doi/10.1093/bjps/axv060/2669734/Evolution-and-Moral-Realism

Strawson, P. F. (2008). *Freedom and Resentment and Other Essays*. Routledge.

Strohminger, N., & Kumar, V. (Eds.). (2018). *The Moral Psychology of Disgust*. Rowman & Littlefield.

Sznycer, D., Al-Shawaf, L., Bereby-Meyer, Y., Curry, O. S., De Smet, D., Ermer, E., . . . Tooby, J. (2017). Cross-cultural regularities in the cognitive architecture of pride. *Proceedings of the National Academy of Sciences*, 114(8), 1874. doi:10.1073/ pnas.1614389114

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual Review of Psychology,* 64(1), 231-255. doi:10.1146/annurev-psych-113011-143812

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1), 35-57. doi:10.1086/406755

Trivigno, F. (2013). The virtue ethical case for pacifism. In M. Austin (ed.), *Virtues in Action*. New York: Palgrave.

Williams, B. (1985). *Ethics and the Limits of Philosophy.* Harvard University Press.

Wong, D. (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.

Wong, D. (2014). Integrating philosophy with anthropology in an approach to morality. *Anthropology Today*, 14(3): 336-55.

Yalçındağ, B., Özkan, T., Cesur, S., Yilmaz, O., Tepe, B., Piyale, Z. E., ... & Sunar, D. (2019). An investigation of moral foundations theory in Turkey using different measures. *Current Psychology*, *38*(2), 440-457.