Explaining Moral Behavior

A Minimal Moral Model

Magda Osman[1] and Alex Wiegmann[2]

[1]Biology and Experimental Psychology Centre, Queen Mary University of London, UK
[2]Cognitive and Decision Sciences, Georg-Elias-Müller Institute of Psychology, University of Göttingen, Germany

Abstract: In this review we make a simple theoretical argument which is that for theory development, computational modeling, and general frameworks for understanding moral psychology researchers should build on domain-general principles from reasoning, judgment, and decision-making research. Our approach is radical with respect to typical models that exist in moral psychology that tend to propose complex innate moral grammars and even evolutionarily guided moral principles. In support of our argument we show that by using a simple value-based decision model we can capture a range of core moral behaviors. Crucially, the argument we propose is that moral situations *per se* do not require anything specialized or different from other situations in which we have to make decisions, inferences, and judgments in order to figure out how to act.

Keywords: morality, decision-making, values, moral philosophy, moral psychology

This special issue includes contributions from both psychologists and philosophers, and various shades in between. In recognition of the multidisciplinary approach taken here, we begin by exploring the relationship between (moral) psychology and (moral) philosophy, and consider whether empirical findings from one field carry any significant implications for the other, and whether both fields can mutually benefit each other. In the second part of this piece we briefly discuss proposals that purport moral behavior to be unique, after which we then consider in more detail the alternative proposal that moral behavior is in fact supported by domain-general processes. In the final section we show just how a domain-general model of decision-making can easily be adapted to explain moral behavior, resulting in our minimal model of morality (henceforth referred to as MMM).

The Relationship Between Moral Psychology and Moral Philosophy

Recently there had been a heated exchange in the New York Review of Books between a philosopher (Shaw, 2016) and two psychologists (Haidt & Pinker, 2016). Shaw, the philosopher, argued that the field of moral psychology lacks a moral compass, and should acknowledge that as a field its dependency on psychological and biological facts makes it morally irrelevant, and reveals nothing about moral propositions of a normative nature. The two psychologists, Haidt and Pinker, replied that their research was never supposed to be understood as a normative guide – and that this should not only be obvious but is also explicitly

stated in the works Shaw was attacking. For instance, Haidt pointed to the following passage at the end of his book:

> "Philosophers typically distinguish between descriptive definitions of morality (which simply describe what people happen to think is moral) and normative definitions (which specify what is really and truly right, regardless of what anyone thinks). So far in this book I have been entirely descriptive." (Haidt, 2012, p. 271).

If this represents the field in general, then moral psychology is a descriptive science, focused on investigating the "is" by looking at how people actually behave in moral situations; this includes how they actually judge whether a certain action is permissible, forbidden, or obligatory. Moral philosophy, as a normative science, investigates the "ought" by considering how people ought to behave and which acts ought to be morally permitted, forbidden, or obligatory – whether or not people behave in this prescribed way in their everyday life. Thus, as Hume (1739) had already observed, you cannot readily deduce an "ought" from an "is". And so, no matter how many parents hit their children or how high the percentage of people might be who believe that homosexuality is wrong, these empirical details do not tell us whether the actual corresponding actions and beliefs are right in a normative sense. Similarly, this line of reasoning can be used to cast doubt on inferring "is" from "ought": Whatever moral philosophy states to be the right code of conduct, it does not seem to tell us much about how people *actually* behave.

By characterizing the pursuits of psychology and philosophy in this way it seems hard to imagine how they could benefit from each other, since the pursuit of each does not seem to feasibly have any bearing on the other. We believe, however, that moral psychology and moral philosophy do have a point of connection, and this in turn can help to generate insights in such a way that each discipline can benefit from the other. The point of connection is *moral intuitions*. In moral philosophy intuitions play an important role in both developing and evaluating normative moral principles and theories. The main purpose in moral psychology is to find principles and theories that can explain people's moral intuitions.

In moral philosophy when it comes to developing a normative principle or theory the philosopher can use her own case-based moral intuitions as data points guiding her toward a normative principle or theory (see Kamm, 2007, for a prime example). In turn, when evaluating a normative moral principle/theory a relevant criterion that can be used is the extent to which the principle/theory is applicable to concrete situations that accord with anyone's own moral intuitions about these situations. If a normative moral theory is applied to particular cases and tells us to do something that stands in sharp contrast to what we believe would be the right thing to do, that is with our moral intuitions, then this mismatch counts strongly against this theory (for a discussion on this, see Hooker, 2000).

Having shown the importance of moral intuitions in moral philosophy we can now consider how moral psychology can be helpful for moral philosophy. Since moral psychology investigates, among others, the psychological mechanisms underlying our moral intuitions, psychology can reveal which intuitions are likely to be biased or caused by morally irrelevant factors, and which intuitions we can trust.[1] Philosophers do not want their theories to be based on flawed intuitions. In fact recently, as a means of revealing insights into intuition philosophers have endorsed experimental methods, along the lines of those of psychology (Knobe & Nichols, 2008). This new approach is a branch of philosophy, called "Experimental Philosophy," which has been characterized by its pioneers as such:

> First we use […] experimental results to develop a theory about the underlying psychological processes that generate people's intuitions; then we use our theory about the psychological processes to determine whether or not those intuitions are warranted. (Knobe & Nichols, 2008, p. 8).

This perfectly captures our key point, which is that psychology can benefit philosophy, but can philosophy benefit psychology? Since moral intuitions play an important role in developing normative principles and theories, moral psychology can use these normative theories and principles as promising starting points in the search for an adequate descriptive moral theory. For example, a key component in a leading descriptive moral theory, such as the Universal Moral Grammar theory (Mikhail, 2011), is the principle of double effect, a principle that has been – and still is – very prominent in moral philosophy (cf. Kamm, 2007). Roughly, it states that intending bad outcomes, in contrast to only foreseeing them, is wrong. This established normative moral principle is used to account for a quite popular phenomenon in moral psychology (cf. Greene, 2014), the so-called trolley problem. This "problem" refers to the question of why most people think that it is permissible to redirect an out-of-control train from a track with five persons (who would otherwise die) onto a track on which it runs over one person (who would otherwise live). Curiously most people think that it would be wrong to push a heavy man from a bridge into the path of the moving train in order to stop it, even if this is the only possible means of saving the five lives. While in the first case (labeled "Switch") it can be plausibly claimed that there was no intention to kill the one person (if she can leave the tracks in time one would be glad) but only to save five people. It isn't possible to plausibly claim in the case labeled "Push" that the action was not based on the intention to harm the heavy person because he is the necessary means to save the five people (if the heavy man runs away one would not be able to save the five lives anymore). In the many demonstrations of the principle of double effect, few participants are able to express the principle when asked to give explanations for why their moral judgments differ in two superficially similar looking scenarios. Clearly this shows two things, first, that the principle is commonly found in laypeople's moral judgments, and second, that the principle was derived from normative investigations in philosophy and not empirical work in psychology. Thus, this provides a good example of how moral philosophy can indeed complement, support, and benefit empirical lines of research in moral psychology.

To sum up, we believe that moral philosophy can benefit from moral psychology by gaining knowledge about which intuitions are prone to biases and which aren't. In turn, moral psychology can benefit from moral philosophy by taking normative moral theories and principles as potential

---

[1] Of course, the question of which intuitions we can trust is again a normative question and since there is – up to now – no agreement about which normative moral theory is the right one it might seem as if moral psychology cannot help but make progress. However, even in moral philosophy there is sometimes agreement. For instance, everyone agrees that the order of presentation of moral dilemmas should not influence our moral intuitions about them. Hence, research that shows that some dilemmas are prone to order effects might tell us that we should not trust our intuitions (for research about order effects see Osman, 2015a, 2015b; Wiegmann & Waldmann, 2014).

candidates for a reasonable descriptive theory of moral behavior. Needless to say, we strongly endorse interdisciplinary research. We have presented the position that moral intuitions are the point of connection between moral psychology and moral philosophy. Therefore, based on our position, moral psychology explores moral intuitions, and the theoretical approaches that are taken to explain them, which have implications for moral philosophy. A common position that impacts both theory and empirical practice is that moral intuitions are special and should be treated as psychologically different to other intuitions that shape and influence our actions. The aim of the next section is to interrogate this idea by briefly considering what the claims are for why moral behavior is claimed to be special, and then examining the extent to which moral behaviors are examples of domain-general rather than domain-specific psychological processes.

Are Moral Intuitions Special?

On the one hand, there are approaches that view moral intuitions such that people consciously deliberate over the moral situation until they come to a decision, judgment, reason, or course of action (Kennett & Fine, 2009; Kohlberg, 1969; Kohlberg, Levine, & Hewer, 1983; Piaget, 1932/1965; Rest, 1983). On the other hand, there are also approaches that view moral intuitions as fundamentally driven by nonconscious processes that in turn are guided by emotions, or formed of them (Ditto, Pizarro, & Tannenbaum, 2009; Greene & Haidt, 2002; Haidt, 2001, 2003; Kagan & Mount, 1981).

In the latter case, to account for evidence that moral intuitions are both driven by nonconscious processes and emotions, a highly popular proposal is that there is a module, which is innate. It has been selected for the purposes of processing information specifically concerning moral situations, and in turn generates specific responses to those situations (e.g., Greene, 2014; Greene & Haidt, 2002; Haidt, 2001; Haidt & Joseph, 2008; Hauser, 2010; Lieberman, 2013; Mikhail, 2011). This is based on a conceptualization of both a modularized mind which is supported by evolutionary adaptive mechanisms. This kind of position earmarks moral behavior as having a unique and privileged status relative to other nonmoral behaviors.

Such strong proposals have been hotly debated, and concerns have been raised around whether there is convincing support regarding the special (e.g., innate, universal, possessing unique features, etc.) nature of moral intuitions (cf. Waldmann et al., 2012). It would be beyond the scope of this paper to evaluate all arguments that have been made for and against the uniqueness of moral behavior. Thus far it is still unclear as to whether moral behavior is unique. So, we consider this sufficient grounds to explore the

alternative, which is that moral intuitions and other types of moral behaviors are underpinned by general psychological processes found in other areas of cognition. It is on this fundamental view that we build our claims for proposing our minimal model of morality (MMM). The MMM assumes that moral situations per se do not invoke anything that is specialized or different from other situations in which we make decisions, inferences, and judgments.

Domain General Then?

If moral behavior is based on domain-general processes (rather than on processes that are unique to the moral domain), then the first step is to specify exactly what it is that constitutes domain generality. A domain-general *lite* proposal would entail the following. The inputs from moral contexts (e.g., moral actions, moral outcomes) are domain specific and are represented in a different manner to inputs from other contexts, but the processes that operate on the inputs are themselves domain general. That is, the operations of reasoning, decision-making, judgment processes are not uniquely different in moral contexts to nonmoral contexts. A domain-general *strong* proposal would be that the inputs from moral contexts are represented in the exact same way as nonmoral contexts, and the processes that operate on them are domain general. Having now set out these different versions of domain generality, the next thing to do is to consider whether there is in fact evidence for the stronger version of the two.

Evidence for Domain-General Inputs in Moral Contexts

One avenue that has been used to explore the extent to which the inputs from moral contexts are similar to those of nonmoral contexts is to look at the framing of a moral action and see whether this influences moral judgments in a pattern similar to nonmoral judgments. The rationale here is the following: If the moral action is couched in such a way that it alerts people to factors that are of relevance when making judgments, inferences, and decisions for nonmoral actions, then if the representational system treats moral and nonmoral actions the same, manipulating the nonmoral factors in a moral scenario will exhibit the same influence on the evaluation of the moral action as it does for nonmoral actions. Spranca, Minsk, and Baron (1991) varied the framing of an action (commission/omission) while holding constant all other details of a moral dilemma with respect to intentions, motives, and consequences. Participants were required to make moral judgments, causal attributions, and punishment decisions regarding

several moral scenarios. For instance in one of the experiments the protagonist in the scenario is privy to information that could sabotage the chances of their tennis opponent winning in the finals of a tennis tournament. In another scenario a protagonist is an eye witness to a car crash in which their testimony could implicate an individual they don't like. In both scenarios the protagonist can either act deliberately to cause harm (commission), or fail to act (omission), which in both cases would ultimately result in a harmful outcome. Across several experiments participants consistently rated commissions as worse than omissions. Commissions were associated with higher penalties and also judged as more intentional than omissions. Spranca et al.'s (1991) explanation for the pattern of judgments was in line with general judgment effects such as omission bias (Sugarman, 1989), and regret (Kahneman & Tversky, 1982) in which the expectancy of bringing about an outcome through action rather than inaction is treated as more causally influential. Moreover, the same judgment pattern is in line with general preferences for, and favorable judgments of inaction over action in judgment and decision-making literature, variably referred to as: status quo bias (Samuelson & Zeckhauser, 1988), omission bias (Ritov & Baron, 1992), inaction inertia (Tykocinski, Pittman, & Tuttle, 1995), and choice deferral (Dhar, 1996, for review, see Anderson, 2003).

Cushman and Young (2011) were able to show that peoples' moral judgments and causal attributions are sensitive to factors such as commission/omission along similar lines to Spranca et al.'s (1991) study. However, in their case they did find differences when comparing moral against nonmoral contexts; this is because in their study they directly compared judgments in moral scenarios and their nonmoral analogs, where as in Spranca et al.'s study, only moral scenarios were used. They also manipulated other general causal input factors such as the framing of scenarios with regard to means/side effects (i.e., judgements are harsher for situations when an action is designed to produce a specific outcome, as compared to when the outcome is an foreseen side effect of an action), contact/noncontact (i.e., judgments are harsher for an outcome resulting from direct physical contact than noncontact). These factors generated similar pattern of judgments when presented in nonmoral contexts as well as moral contexts. On this basis Cushman and Young suggested that generally the relative influence of different input factors on judgments in moral and nonmoral contexts was a difference in degree than kind, but that the representational system used to code moral and nonmoral contexts is broadly the same.

Shenhav and Greene (2010) used the same basic rational as Spranca et al. (1991) and Cushman and Young (2011) in order to examine the neurological processes underpinning moral judgment. To do this they investigated the neural activity when judgments were made about moral dilemmas for which the outcomes varied according to factors typically examined in economic domains (e.g., magnitude, probability). Participants were presented with five different moral scenarios, and in each an action was described (e.g., redirecting a speed boat from causing fatal injury). For each scenario ten subsequent trials were presented in which the magnitude of the outcome of the action (e.g., the number of lives saved/lives lost) and the probability of the outcome of the action (e.g., the probability of the lives saved/lives lost) were systematically varied. On each trial participants were required to make a judgment on a 5-point scale (1 = *completely unacceptable* – 5 = *completely acceptable*). The key message from the neuroimaging findings was that the patterns of activation were not unique to moral contexts. Rather, when factors in moral scenarios such as gains versus losses, magnitude, and probability of outcome were manipulated, the pattern of activation was in line with typical economic tasks that solicit expected-value judgments of nonmoral contexts (e.g., ventromedial prefrontal cortex, medial orbitofrontal cortex, central insula, dorsal striatum, and anterior and posterior cingulate cortices; these areas are sensitive to the magnitude of potential loss/gain of choice alternatives).

While not an exhaustive list, taken together with work examining manipulations of domain-general inputs on judgment, decision-making, and choice behavior in moral and nonmoral contexts, empirical work reliably shows that for both, people are sensitive to framing effects (gains/losses, Kern & Chugh, 2009; Petrinovich & O'Neill, 1996; Shenhav & Greene, 2010), omission bias (Bartels & Medin, 2007; Bostyn & Roets, 2016; Cushman & Young, 2011; Cushman, Young, & Hauser, 2006; Gino, Shu, & Bazerman, 2010; Powell, Derbyshire, & Guttentag, 2012; Ritov & Baron, 1999; Spranca et al., 1991), the magnitude of the outcomes of actions (Bartels & Medin, 2007; Connolly & Reb, 2003; Morris & McDonald, 1995; Rai & Holyoak, 2010; Ritov & Baron, 1999; Shenhav & Greene, 2010; Singhapakdi, Vitell, & Kraft, 1996; Stein & Ahmad, 2009; Weber, 1996), and probability (Brand & Oaksford, 2015; Shenhav & Greene, 2010; Singhapakdi et al., 1996; Stein & Ahmad, 2009). The conclusions from these findings are not of course unanimous. Some claim that this evidence supports the fact that moral inputs generate effects that are the same as nonmoral inputs (Rai & Holyoak, 2010), whereas others claim the evidence supports the fact that moral inputs in principle share similar characteristics as nonmoral inputs (Cushman & Young, 2011). However, the overlap between the patterns of behavior at the input level for moral and nonmoral contexts seems considerable.

Evidence for Domain-General Processes in Moral Contexts

The research strategies for examining the possibility that domain-general processes support moral behavior can be grouped into three categories: Category (1) working showing that factors predicting developmental changes in nonmoral contexts are the same as moral contexts; Category (2) manipulating high order cognition (e.g., working memory load, levels of construal, incentives) impacts behavior in the same way in nonmoral and moral contexts; Category (3) general task manipulations in moral contexts and nonmoral contexts reveal similar patterns in judgments, inferences, and decisions. For Categories 2 and 3 the evidence is particularly compelling given that the manipulations in many cases impact behavior despite the fact that they should not have a bearing on moral contexts, we consider this point in more detail later in this section.

Category 1: A recently popular method of assessing moral intuitions in children (aged 6–12 years) involves presenting them with a range of moral dilemmas in pictorial form, in which the child takes on a first person perspective, and observes the depiction of social scenes in which a conflict emerges, and the child is asked to make a decision and then justify it (e.g., observing a scene in a classroom in which they could either choose between: (a) tell the teacher that you saw a classmate cheating on an exam, (b) not say anything about what you saw, Beauchamp & Dooley, 2012; Beauchamp, Dooley, & Anderson, 2013; Vera-Estay, Seni, Champagne, & Beauchamp, 2016). Here choice and justifications of choices are scored separately. The pattern of behavior shifts with age from more egocentrically oriented decisions and justifications (aged 6–8 years), to sensitivity of social factors such as fairness and reciprocity (aged 8–9 years), to an eventual appreciation of compliance to social norms and rules, and the fundamental rights of individuals (aged 10–12 years). Factors that influence developmental changes in moral intuitions are not exclusively predicted by deontic and social knowledge of fairness, harm, justice, welfare, rights of others; some of which have been speculated to be directly associated with the five innate moral foundations proposed by Haidt (2001). Moreover, different moral intuitions reported at different ages are not exclusively predicted by the development of particular emotions such as empathy (Beauchamp et al., 2013; Hoffman, 2000; Miller, Eisenberg, Fabes, & Shell, 1996). Instead, cognitive development seems to most closely track the developmental changes in moral intuitions (Kerr & Zelazo, 2004; Stewart & Pascual-Leone, 1992; Vera-Estay, Dooley, & Beauchamp, 2014; Vera-Estay et al., 2016); this includes working memory, verbal fluency, inhibition, and attentional control. More to the point these same executive functions are thought to mediate high order

cognition (i.e., judgment, reasoning, decision-making, problem-solving) in nonmoral social contexts (Kochanska, Murray, & Harlan, 2000), theory of mind abilities (Devine & Hughes, 2014), and social adjustment (Martel et al., 2007). Similarly, aging research has shown that declining executive functions also track declines in theory of mind ability in nonmoral social contexts (McKinnon & Moscovitch, 2007 –young [approx. age 20] vs. older [approx. age 75] adults; Pratt, Diessner, Pratt, Hunsberger, & Pancer, 1996 – middle-aged [aged 35–54] vs. older [aged 64–80] adults) and in moral contexts (Pratt et al., 1996). Thus, work in the developmental psychology domain and adult aging domain both shows that the factors that influence decision-making in moral and nonmoral contexts overlap to a large degree, and are underpinned by high order executive functions, and not innate moral modules guided by emotions.

Category 2: As well as showing that executive functioning is associated with reasoning and decision-making in moral and nonmoral contexts (McKinnon & Moscovitch, 2007; Moore, Clark, & Kane, 2008; Stewart & Pascual-Leone, 1992), the introduction of various external manipulations that directly effect executive function also affects decision-making in moral and nonmoral contexts. Introducing a cognitive load task is an effective means of burdening executive functions. In moral tasks this entails having the participant make moral judgments while at the same time counting backwards in 3 s, or tone counting (Conway & Gawronski, 2013; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). Alternatively, executive functions can be loaded by severely limiting the time that is available to make a judgment, decision, or inference on a moral or nonmoral task (Körner & Volk, 2014). Some have argued that the impact of cognitive loading and deadlines specifically affects certain types of responses (e.g., utilitarian vs. deontological) that are only found in nonmoral contexts (Conway & Gawronski, 2013; Greene et al., 2008). However, the consistently main pattern of findings is that in general analytical strategies typically employed in moral and nonmoral contexts are severely hampered when cognitive resources become limited, either through cognitive load, or response deadlines (Koop, 2013; Körner & Volk, 2014; McGuire, Langdon, Coltheart, & Mackenzie, 2009; Trémolière, De Neys, & Bonnefon, in press). These findings also challenge claims that suggest that moral judgments and decisions are predominately nonconscious (Haidt, 2001), because if they were, they would be immune to effects of cognitive loading.

Category 3: Several manipulations have been introduced that expose the similarity of behavior in moral and nonmoral contexts which should be superfluous to moral contexts. For instance, Rai and Holyoak (2010) tested principles that were not derived from moral philosophy,

but rather from findings in consumer choice theory and risk perception. When participants were asked to provide more reasons for their decisions (i.e., 2 vs. 7 reasons), participants were less likely to choose an action that involved sacrificing one life to save others. Moreover, the level of agreement between the reasons given was greater when fewer choices were made, a pattern typical of consumer choice behavior. Similarly, other evidence suggests that just as in nonmoral contexts (including management, economic, consumer contexts), moral judgment, reasoning, and decision-making behavior are affected by level of construal (i.e., the relative psychological distance a participant has to the key constructs of a task, Eyal, Liberman, & Trope, 2008; Körner & Volk, 2014), acts of *joint* versus *separate evaluation* (i.e., where two options are evaluated side by side, or separately) (Bartels, 2008; Lombrozo, 2009; Paharia, Kassam, Greene, & Bazerman, 2009), *processing fluency* (i.e., the ease with which information is either accessed or processed, Laham, Alter, & Goodwin, 2009), individual differences in *need for cognition* (Bartels, 2008), order effects (Osman, 2015a, 2015b; Petrinovich & O'Neill, 1996; Wiegmann, Okan, & Nagel, 2012), and *sensitivity to financial rewards* (FeldmanHall et al., 2012). Crucially, moral contexts are unique, and the behavior that they elicit is special, then they should not be consistently subject to the same factors that impact judgment, decision-making, and reasoning across such a variety of nonmoral domains.

As an alternative to manipulating general task instructions, there is work examining the way in which directly manipulating causal structures reveal patterns of behavior that are common to both moral and nonmoral contexts (for review see, Alicke, Mandel, Hilton, Gerstenberg, & Lagnado, 2015). For instance, Iliev, Sachdeva, and Medin (2012) presented participants with short videos of objects (cylinders and cones, and fireballs that were "harmless" to the cylinders but lethal to the cones) that were in motion. The interactions between objects were manipulated so that an object (e.g., cylinder) could either be static or moving, and it could causally affect another object (either cone, fireball), for which the consequences of that interaction could differ (cone/fireball remains static or moving). In addition, the frequency of the interactions, as well as the force with which collisions between objects occurred was also manipulated. The order in which objects interacted with each other, the cause-effect relationships between objects, as well as the causal intensity of the interaction (force was used as a proxy for this), contributed to judgments of "blame" of objects intending to "harm" other objects. Crucially, the blame judgments were compared to an analog of the same interactions, but couched in short vignettes of social interactions based around the classic Trolley Problem. The findings revealed that the pattern of judgments for abstract stimuli in which the salient features were

the causal relations rather than any moral or social interactions, generated the same pattern of judgments when the social features were made salient. Moreover, the strong role of causal attributions has been shown to account for patterns in moral judgments and decisions (Bostyn & Roets, 2016; Lagnado & Channon, 2008), and when causal features of the moral scenario are manipulated (Iliev et al., 2012; Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Wiegmann & Waldmann, 2014).

Taken together, we would argue that the work examining domain-general factors associated with moral behavior provides plausible support for an alternative to the popular domain-specific models of moral intuitions. We use this as a platform for adopting the domain-general strong proposal from which the aim in the next section is to present our alternative minimal model of moral behavior.

A Minimal Model of Moral Behavior (MMM)

Value-based decision-making is a general psychological term used to describe decision-making in situations in which we face a choice between options associated with different rewards (e.g., monetary, social) and costs (e.g., effort, risk, outcome probability). The moral psychology domain has often made reference to value-based decision-making models for the reason that parallels can be drawn between economic and moral decision-making which can inspire experimental manipulations (Bartels & Medin, 2007; Caspers et al., 2011; Crockett, Kurth-Nelson, Siegel, Dayan, & Dolan, 2014; Crockett et al., 2015; FeldmanHall et al., 2012; Rai & Holyoak, 2010; Shenhav & Greene, 2010; Tassy et al., 2011). But also, there is a rationale to connect value-based decision-making to the moral domain because it has been speculated that the very same neural circuitry underpins decision-making under risk and uncertainty in nonmoral contexts and moral contexts. In fact, many moral psychologists (Crockett et al., 2014, 2015; Han, Chen, Jeong, & Glover, 2016; Shenhav & Greene, 2010, 2014; Vartanian, Mandel, & Duncan, 2011) have directly drawn from current neuroeconomic models of value-based decision-making (particularly Rangel, Camerer, & Montague, 2008) as a way to understand the psychological and neurological processes involved in subjective valuation of options in moral contexts.

Thus, moral psychologists have borrowed heavily from value-based decision-making models, particularly neuroeconomic ones (Rangel et al., 2008) that offer an opportunity to model decision-making processes. One might ask, why not simply use traditional judgment and decision-making models such as Prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992)? In actual fact, it isn't a situation of either/or, since several neuroeconomic models of value-based decision-making (Assadi, Yucel, & Pantelis, 2009; Doya, 2008; Ernst & Paulus, 2005; Kable

Table 1. Summary of stages included in different value-based decision-making models. The shaded cells denote theories that make reference to a particular decision-making stage

Decision-making stage	Current value-based decision-making models					
	Assadi et al. (2009)	Doya (2008)	Ernst & Paulus (2005)	Kable & Glimcher (2009)	Rangel et al. (2008)	Rigoux & Guigon (2012)
Representation		▓	▓	▓	▓	
Valuation	▓	▓	▓	▓	▓	▓
Action selection	▓	▓	▓	▓	▓	▓
Action execution	▓		▓			▓
Outcome evaluation		▓			▓	
Learning					▓	

& Glimcher, 2009; Rangel et al., 2008; Rigoux & Guigon, 2012) are built on the fundamentals of key models of risk and uncertainty, such as Prospect theory, particularly when it comes to ascribing what goes on in the valuation stage of the decision-making process. While there is an interest in extending aspects of value-based decision-making models to moral psychology, there has been no attempt to explore the possibility that value-based models in and of themselves are sufficient to describe the processes involved in moral decisions/judgments and reasons. In this section, we outline the conceptual apparatus of a value-based decision-making model that is context-free, and show that, it can very easily be extended to explain choice behavior in moral decision-making contexts. To do this, we take the approach that domain-general processes, rather than domain-specific ones, support moral behaviors. The aim of this last section then is to condense value-based decision-making models (Assadi et al., 2009; Doya, 2008; Ernst & Paulus, 2005; Kable & Glimcher, 2009; Rangel et al., 2008; Rigoux & Guigon, 2012), into their essential features (i.e., stages of the process and the components of each stage) to form our MMM, in abstract form first. From which we then show how the same features easily extend to moral contexts. This discussion should serve as, and in principle, proof of concept that a value-based model in and of itself can be used to account for moral behavior.

Neuroeconomic Value-Based Decision-Making Models

The most comprehensive model of value-based decision-making has been proposed by Rangel et al. (2008) in which there are five consecutive stages: (1) *Representation*, (2) *Valuation*, (3) *Action Selection*, (4) *Outcome Evaluation*, and (5) *Learning*. The Representation stage involves an analysis of the external state of the environment, the internal state of the decision-maker, and the attributes of the available courses of action (see also: Regan, 2014). During the Valuation stage decision-makers integrate different dimensions (probability, effort, rewards, time) of each available

option into a subjective value (in line with Prospect theory). This subjective value is thought to depend on the learned values of actions taken in the past. During the Action Selection stage these subjective values are compared, and the action which leads to an outcome with the highest subjective value is selected, from which a post-outcome evaluation process occurs which assesses the desirability of an outcome. The comparison that takes place at this stage assesses the discrepancy between the estimated and the experienced outcome, which is coded as a prediction error signal (again in line with Prospect theory). As a consequence of Learning, decision-makers update the action-outcome associations and use the error signal as a way to adjust future behaviors to reduce the prediction error (Table 1).

Crucially, Rangel et al. (2008) claim that during the action selection stage choices are made and actions simply follow from the choices made, which means that there is no need for a separate action execution stage. Why would this matter? Several other models suggest otherwise (Assadi et al., 2009; Ernst & Paulus, 2005; Rigoux & Guigon, 2012). These models view *Action Selection* as a stage in which there is an analysis of the costs and benefits associated with different options, and that the outcome of this analysis reveals an overall preference toward one of the options. *Execution*, on the other hand, relies on motivation (i.e., mobilization of energetic resources) and planning of sequences of actions to achieve the goal. Just as one is about to execute an action one may revisit the value attached to the action as the potential costs become more salient. For this reason it is important to disentangle selection from execution. This is because, by implication, an individual might evaluate two options (option A, option B) for which option A is selected, but when it comes to performing the action associated with the option, both options are reevaluated, and the action associated with option B is executed. This distinction is not only critical for value-based decision-making in general, but also in the context of moral behavior (which we will discuss in connection with predictions at the end of this section).

Based on this critical distinction, the most comprehensive framework that integrates the insights from the
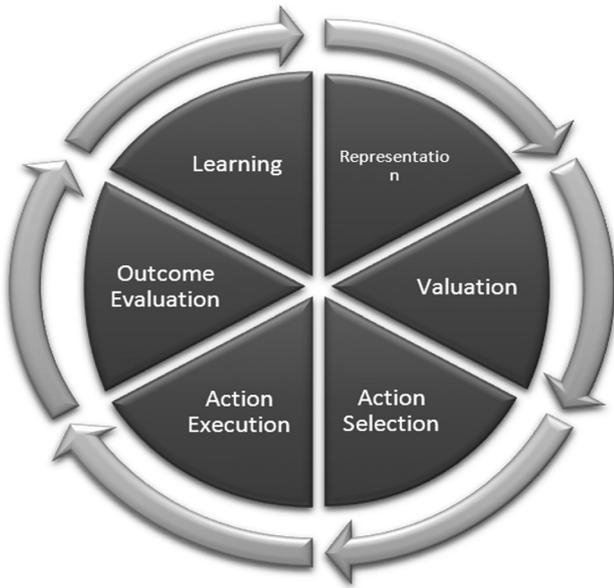
Figure 1. Dynamic-value-effort decision-making model.

reviewed value-based decision-making models is the Dynamic-Value-Effort-based Decision-Making model (D-V-E-D-M model, Ludwiczack, 2016). This framework describes six consecutive and interdependent stages of decision-making: *Representation*, *Valuation*, *Action Selection*, *Action Execution*, *Outcome Evaluation*, and *Learning* (see Figure 1). This forms the basis of our Minimal Moral Model.

Extensions to Our Proposed Minimal Moral Model (MMM)

Here we outline the different stages of the decision-making process of the D-V-E-D-M model in general terms and then specifically in the context of a moral decision-making context. We argue here that the D-V-E-D-M can be repurposed as a *Minimal Moral Model Theory (MMM)* and serves as a psychological model of moral decision-making. The model operates under two basic assumptions:

1. Human beings are good at acquiring norms and implementing norms;
2. The relations between agents and events are interpreted from the view of causal structures.

These assumptions are based on consolidating work from the evidence reviewed in the previous section on the domain generality of moral behavior. It is important to clarify that we are focusing specifically on moral decision-making behavior, and not on moral judgment or moral reasoning, though stages of the process of decision-making will connect with aspects of moral judgment and reasoning; the process of judgment is an estimation/evaluation of a moral situation (e.g., How bad is the outcome? How responsible is the agent

for the outcome?) and moral reasoning involves an inference regarding features of the moral scenario (e.g., What are the agent's intentions? What are the consequences of an action?). What we are specifically concerned with here is detailing the psychological process regarding what an individual does when faced with a set of options that they self-proclaim as resulting from a moral situation; crucially here we are careful not to impose a judgment of what constitutes a moral situation or not, this is self-determined. Another important factor that we clarify here is that the process of decision-making we are describing is one that concerns an individual that decides what course of action to take, and not the process by which collective action is taken in response to a moral dilemma. We remain agnostic with respect to whether the decision-making process we describe here scales up to group level moral decision-making behaviors.

In the *Representation stage* what is encoded is: (1) the potential courses of action that are available in the decision-making scenario, (2) the internal state of the decision-maker, and (3) the external state of the environment. Within the context of a moral scenario various options will become available to the agent. The options that the agent faces will be embedded in a causal structure of the moral scenario, which will be more or less complex depending on how the agent chooses to conceptualize it. There will be emotional and physiological experiences (e.g., hunger, pain) that will either result from or coincide with the decision-making situation the individual faces. Along with this there will be various background details regarding norms (social, cultural, religious, political, legal, organizational, community, familial) and personal practices (prior experiences with similar type scenarios) that the individual may choose to use to inform their characterization of the moral scenario.

There is a large body of work suggesting that emotional states influence the kinds of decisions and judgments people make in moral contexts (Eskine, Kacinik, & Prinz, 2011; Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005), there is also work showing, particularly in the developmental domain, that the depth of knowledge of norms and conventions impacts decisions and judgments in moral contexts (Nichols, 2002; Schmidt, Rakoczy, & Tomasello, 2012; Schmidt & Tomasello, 2012; Sripada & Stich, 2005). Moreover, there is also evidence that different political affiliations influence people's moral judgments and decisions (Graham, Haidt, & Nosek, 2009). Therefore, this work constitutes support for the claim that the inputs that go into the representational stage of moral decision-making include personally or contextually salient features such cultural, emotional, and deontic factors.

Once an individual represents the options the *Valuation stage* appraises the costs (i.e., effort) and benefits (i.e., rewards) associated with each option. When assessing

the benefits associated with different options, factors such as the quality and quantity of a reward (Green & Myerson, 2004; Kacelnik & Bateson, 1997), as well as the valence (positive/negative) and salience (intensity, magnitude) of a reward (Ernst & Paulus, 2005) are taken into account. The costs associated with different options will take into account the effort (physical, mental) in achieving the outcome, the probability of achieving that outcome, and time needed to achieve it. The subjective value of each option is abstract (it is a subjective construct that indicates the general positive or negative property overall of each option) – it is a form of common currency which drives choice during the Action Selection stage (Brosch & Sander, 2013). The subjective value of an option is also thought to be context dependent, that is, the value assignment depends on the external environment (e.g., reliability of reward predicting cues), as well as internal factors (e.g., motivation, emotional and physiological experiences, learning history of similar past situations the decision-maker has faced, Doya, 2008).

The large body of behavioral and neuroimaging work suggests that people are sensitive to the framing of moral actions with respect to valence (Kern & Chugh, 2009; Petrinovich & O'Neill, 1996; Shenhav & Greene, 2010), probability (Brand & Oaksford, 2015; Shenhav & Greene, 2010; Singhapakdi et al., 1996; Stein & Ahmad, 2009), and magnitude of the outcome (Bartels & Medin, 2007; Connolly & Reb, 2003; Morris & McDonald, 1995; Rai & Holyoak, 2010; Ritov & Baron, 1999; Shenhav & Greene, 2010; Singhapakdi et al., 1996; Stein & Ahmad, 2009; Weber, 1996) when making decisions and judgments in moral and nonmoral contexts. This supports the view that these dimensions do indeed inform the valuation of choice alternatives generally, and those faced in moral contexts.

Action Selection and Action Execution: The *Action Selection* relies on subjective values of options computed during the Valuation stage. During Action Selection these subjective values are compared and the option with the highest value is chosen (Koopmans, 1960; Neumann & Morgenstern, 1944; Prelec & Loewenstein, 1991; Rachlin, Battalio, Kagel, & Green, 1981). During *Action Execution* the preferences of the decision-maker expressed through choice are translated into actual actions. Crucially, execution of an action is not exclusively driven by rewards per se (Brehm & Self, 1989). This claim takes into account phenomena in which initial valuations support choice for a particular action, but that when it comes to executing the action, further elaboration of the choice alternatives ultimately leads to a different action being executed than had been originally selected.

There is justification for separating out action selection from action execution, because similar distinctions have been proposed in the moral domain, namely the judgment-action gap (Jennings, Mitchell, & Hannah, 2015;

Walker, 2004). For instance, FeldmanHall et al., (2012) provide good evidence of the gap between what people say they would do when faced with a moral decision, and what they actually do when performing an action that has real consequences. That is to say, at the stage of choice, what we want to do may be consistent with good intentions that do not get realized in practice, particularly when monetary incentives are presented that create a conflict of interest. Moreover, there is work suggesting that prompting further elaboration of the same initial judgment/choice can lead to differences in actual judgments/choices (Rai & Holyoak, 2010). Taken together this lends support to the claim that further elaboration of one's processing of the options in moral context leads to changes in the valuation of the option such that actions selected in association with a particular option end up changing when faced with the reality of executing those actions.

Outcome Evaluation and Learning: During the *Outcome Evaluation* stage desirability of the experienced outcome is assessed and compared to the estimations made during the Valuation stage, and this is based on a relative assessment of the values rather than on absolute terms. The discrepancy between the predicted and experienced outcome (i.e., prediction error) is thought to serve as a basis for updating the action-outcome associations represented neurologically and psychologically, which in turn triggers the *Learning* stage (Rangel et al., 2008). Reward feedback in particular is one of main factors taken into account during the Outcome Evaluation and Learning stages of the model.

Several recent studies (Osman, 2015a, 2015b) suggest that when presented with moral scenarios in which different critical details of the scenario are presented at different stages, people adjust and update their consideration of the moral scenario accordingly. Osman (2015a, 2015b) presented participants with a news story comprised of a headline, main story, and conclusion. After each section of the story participants made judgments and decision regarding whether the actions of the protagonist were morally permissible or not. In both studies, the findings show that participants are sensitive to the information presented at each stage, and learn, by updating and adjusting their behavior accordingly. This is consistent with similar findings revealing updating of moral judgments and beliefs when new information is presented (Horne, Powell, & Hummel, 2015; Horne, Powell, & Spino, 2013).

Based on the MMM that we introduced, it is possible to generate several new testable predictions. For instance, when making moral decisions, during the action selection and execution stages, do people pay attention to the magnitude of the outcome on relative or absolute terms? The model would predict that the basis on which an action selection is made, will be on relative terms. More to the point, the valuation of options at the action execution stage

of taking a moral action will differ from that of the action selection stage. At the action execution stage, what becomes salient is the difficulty (indexed through measures of effort) in implementing the various moral actions available in the context, and it is for this reason that the valuation of options differs from the action selection stage. Put simply, the relative assessment of rewards will influence action selection, but other factors such as difficulty in actually carrying out the action will take greater precedence at the action execution stage, which is why the valuation of options differs at these two stages of the decision-making process. Moreover, we also predict that emotion experiences will track changes in the valuation of options at the selection stage and execution stage, for the same reasons that other properties of the options become reevaluated. Which means that emotional states are dynamic and should correspond with the shifting valuation of options at different stages of the decision-making process.

In addition, the temporal relationship between action selection and action execution will affect the valuation of options at the action execution stage, but not at the action selection stage. Why? Because the point at which the selection of actions is made are prospects until the point of execution. This means that nothing will change with regard to the valuation of options regardless of whether the selected action is going to be implemented minutes, hours, or days after the action selection is made. If the execution of the action directly follows selection of the action, then there is less time to carry out a reevaluation of the options. If the execution of the action is delayed in time, then the longer the interval between selection and execution, the greater the opportunity to reevaluate the options, and reconsider taking the action originally selected. These differences help to also explain the basis on which the judgment-action gap occurs, and further elaborate characteristic differences in the decision-making process that underpins this phenomenon, and predict where they are likely to be detected.

To sum up, based on the work we reviewed, we have proposed a domain-general approach to moral behavior as a promising direction to take and introduced a model with testable hypothesis derived from it. In addition, we have shown that, in principle, a domain-general model can be modified in order to account for moral behavior, with supporting evidence from the moral domain.

Alpha and Omega: Implications for Moral Psychology and Moral Philosophy

Let us go back to the starting point of this paper where we asked whether findings in moral psychology have implications for the field of moral psychology – and vice versa. Let us also assume that the claims and the predictions we laid out in the previous section turn out to be well supported. What would be the implications for moral psychology and moral philosophy?

The most important implication for the field of moral psychology would be that it is possible to build on domain-general principles of moral behavior from existing reasoning, judgment, and decision-making research. This in turn means enormous progress for the field of moral psychology because it can defer to theories in other areas of psychology – for example, theories in decision-making, judgment, and reasoning – which are currently much more advanced. Moreover, instead of developing new principles or theories for a phenomenon encountered in a moral setting, researchers in moral psychology could consult the existing literature where an explanation of their finding is likely to already exist. To illustrate this point consider a study that examines moral intuitions in different groups. Wiegmann, Horvath, and Meyer (under review) show that laypeople's and philosopher's moral intuitions around the trolley problem vary as a function of the number of available alternative options presented to them. Assuming that moral judgments are somehow special would in turn imply that these researchers would have to develop, from scratch, a theory of this based on unique preference reversal behavior specifically found in moral contexts. If instead moral behavior is viewed as a domain-general process, the findings can easily be accounted for based on existing literature from judgment and decision-making research such as Tversky's (1969) work on intransitive preferences.

The same benefits of this research approach extend to the moral philosophy domain. As we described at the beginning of the paper, empirical research can inform philosophers as to which moral intuitions are likely to be biased. If moral judgments, decisions, and behavior can be captured by well-developed domain-general theories then our theoretical and empirical resources for gaining knowledge about moral intuitions would be much greater, as compared to the recourses provided by moral psychology alone. As in the example above, instead of waiting for moral theories coming from moral psychology, philosophers could already start to study the relevant existing psychological literature to help inform theorizing around moral intuitions.

In conclusion, acknowledging the possibility that moral judgments, decisions, and behavior might be explained by domain-general processes and theories could enormously advance both moral psychology and moral philosophy. Standing on the shoulders of giants of established theories and models would let us see much further than where we currently stand.

References

Alicke, M. D., Mandel, D. R., Hilton, D. J., Gerstenberg, T., & Lagnado, D. A. (2015). Causal conceptions in social explanation and moral evaluation. A historical tour. *Perspectives on Psychological Science, 10*, 790–812. doi: 10.1177/1745691615601888

Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin, 129*, 139–167. doi: 10.1037/0033-2909.129.1.139

Assadi, S. M., Yucel, M., & Pantelis, C. (2009). Dopamine modulates neural networks involved in effort-based decision-making. *Neuroscience and Biobehavioral Reviews, 33*, 383–393. doi: 10.1016/j.neubiorev.2008.10.010

Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition, 108*, 381–417. doi: 10.1016/j.cognition.2008.03.001

Bartels, D. M., & Medin, D. L. (2007). Are morally motivated decision makers insensitive to the consequences of their choices? *Psychological Science, 18*, 24–28. doi: 10.1111/j.1467-9280.2007.01843x

Beauchamp, M. H., & Dooley, J. J. (2012). *Administration and Coding Manual Socio-Moral Reasoning Aptitude Level Task (SoMoral)*. Montreal, QC: ABCs Laboratory.

Beauchamp, M. H., Dooley, J. J., & Anderson, V. (2013). A preliminary investigation of moral reasoning and empathy after traumatic brain injury in adolescents. *Brain Injury, 27*, 896–902. doi: 10.3109/02699052.2013.775486

Bostyn, D. H., & Roets, A. (2016). The morality of action: The asymmetry between judgments of praise and blame in the action – omission effect. *Journal of Experimental Social Psychology, 63*, 19–25. doi: 10.1016/j.jesp.2015.11.005

Brand, C. M., & Oaksford, M. (2015). The effect of probability anchors on moral decision making. *Proceedings of the Cognitive Science Society*. Pasadena, CA.

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology, 40*, 109–131. doi: 10.1146/annurev.ps.40.020189.000545

Brosch, T., & Sander, D. (2013). Neurocognitive mechanisms underlying value-based decision-making: From core values to economic value. *Frontiers in Human Neuroscience, 7*, 398. doi: 10.3389/fnhum.2013.00398

Caspers, S., Heim, S., Lucas, M. G., Stephan, E., Fischer, L., Amunts, K., & Zilles, K. (2011). Moral concepts set decision strategies to abstract values. *PLoS One, 6*, e18451. doi: 10.1371/journal.pone.0018451

Connolly, T., & Reb, J. (2003). Omission bias in vaccination decisions: Where's the "omission"? Where's the "bias"? *Organizational Behavior and Human Decision Processes, 91*, 186–202. doi: 10.1016/S0749-5978(03)00057-8

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology, 104*, 216–235. doi: 10.1037/a0031021

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences, 111*, 17320–17325. doi: 10.1073/pnas.1408988111

Crockett, M. J., Siegel, J. Z., Kurth-Nelson, Z., Ousdal, O. T., Story, G., Frieband, C., & Dolan, R. J. (2015). Dissociable effects of serotonin and dopamine on the valuation of harm in moral decision making. *Current Biology, 25*, 1852–1859. doi: 10.1016/j.cub.2015.05.021

Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science, 35*, 1052–1075. doi: 10.1111/j.1551-6709.2010.01167

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science, 17*, 1082–1089. doi: 10.1111/j.1467-9280.2006.01834x

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development, 85*, 1777–1794. doi: 10.1111/cdev.12237

Dhar, R. (1996). The effect of decision strategy on deciding to defer choice. *Journal of Behavioral Decision Making, 9*, 265–281. doi: 10.1002/(SICI)1099-0771(199612)9:4<265::AID-BDM231>3.3.CO;2-W

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. In D. L. Medin, B. H. Ross, & D. Bartels (Eds.), *Psychology of learning and motivation. Moral judgment and decision making* (Vol. 50, pp. 307–338). Cambridge, MA: Elsevier Textbooks.

Doya, K. (2008). Modulators of decision making. *Nature Neuroscience, 11*, 410–416.

Ernst, M., & Paulus, M. P. (2005). Neurobiology of decision making: A selective review from a neurocognitive and clinical perspective. *Biological Psychiatry, 58*, 597–604. doi: 10.1016/j.biopsych.2005.06.004

Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science, 22*, 295–299. doi: 10.1177/0956797611398497x

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal of Experimental Social Psychology, 44*, 1204–1209. doi: 10.1016/j.jesp.2008.03.012

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition, 123*, 434–441. doi: 10.1016/j.cognition.2012.02.001

Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless + harmless = blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human Decision Processes, 111*, 93–101. doi: 10.1016/j.obhdp.2009.11.001

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029–1046. doi: 10.1037/a0015141

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin, 130*, 769–792. doi: 10.1037/0033-2909.130.5.769

Greene, J. (2014). *Moral tribes: Emotion, reason and the gap between us and them*. New York, NY: Atlantic Books Ltd.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences, 6*, 517–523. doi: 10.1016/S1364-6613(02)02011-9

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*, 1144–1154. doi: 10.1016/j.cognition.2007.11.004

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108. doi: 10.1126/science.1062872

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834. doi: 10.1037/0033-295X.108.4.814

Haidt, J. (2003). The emotional dog does learn new tricks: A reply to Pizarro and Bloom (2003). *Psychological Review, 110*, 197–198. doi: 10.1037/0033-295X.110.1.197

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Vintage.

Haidt, J., & Joseph, C. (2008). 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., & Pinker, S. (2016). Moral psychology: An exchange. *The New York Review of Books*. Retrieved from: http://www.nybooks.com/articles/2016/04/07/moral-psychology-an-exchange/

Han, H., Chen, J., Jeong, C., & Glover, G. H. (2016). Influence of the cortical midline structures on moral emotion and motivation in moral decision-making. *Behavioural Brain Research, 302*, 237–251. doi: 10.1016/j.bbr.2016.01.001

Hauser, M. D. (2010). Moral minds: How nature designed our universal sense of right and wrong. *Journal of Bioeconomics, 12*, 173–177.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press.

Hooker, B. (2000). *Ideal code, real world: A rule-consequentialist theory of morality*. Oxford, UK: Clarendon Press.

Horne, Z., Powell, D., & Hummel, J. (2015). A single counterexample leads to moral belief revision. *Cognitive Science, 39*, 1950–1964. doi: 10.1111/cogs.12223

Horne, Z., Powell, D., & Spino, J. (2013). Belief updating in moral dilemmas. *Review of Philosophy and Psychology, 4*, 705–714. doi: 10.1007/s13164-013-0159-y

Hume, D. (1739). *A treatise of human nature*. London, UK: Oxford University Press.

Iliev, R. I., Sachdeva, S., & Medin, D. L. (2012). Moral kinematics: The role of physical factors in moral judgments. *Memory & Cognition, 40*, 1387–1401. doi: 10.3758/s13421-012-0217-1

Jennings, P. L., Mitchell, M. S., & Hannah, S. T. (2015). The moral self: A review and integration of the literature. *Journal of Organizational Behavior, 36*, S104–S168. doi: 10.1002/job.1919

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron, 63*, 733–745. doi: 10.1016/j.neuron.2009.09.003

Kacelnik, A., & Bateson, M. (1997). Risk-sensitivity: Crossroads for theories of decision-making. *Trends in Cognitive Sciences, 1*, 304–309. doi: 10.1016/S1364-6613(97)01093-0

Kagan, J., & Mount, R. (1981). *The second year: The emergence of self-awareness*. Cambridge, MA: Harvard University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*, 263–291. doi: 10.2307/1914185

Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition, 11*, 143–157. doi: 10.1016/0010-0277(82)90023-3

Kamm, F. M. (2007). *Intricate ethics*. Oxford, UK: Oxford University Press.

Kennett, J., & Fine, C. (2009). Will the real moral judgment please stand up? *Ethical Theory and Moral Practice, 12*, 77–96. doi: 10.1007/s10677-008-9136-4

Kern, M. C., & Chugh, D. (2009). Bounded ethicality: The perils of loss framing. *Psychological Science, 20*, 378–384. doi: 10.1111/j.1467-9280.2009.02296x

Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The children's gambling task. *Brain and Cognition, 55*, 148–157. doi: 10.1016/S0278-2626(03)00275-6

Knobe, J. & Nichols, S. (Eds.). (2008). *Experimental philosophy*. Oxford, UK: Oxford University Press.

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*, 220–232. doi: 10.1037/0012-1649.36.2.220

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago, IL: Rand McNally.

Kohlberg, L., Levine, C., & Hewer, A. (1983). Moral stages: A current formulation and a response to critics. In J. A. Meacham (Ed.), *Contributions to human development, Vol. 10.* Basel, Switzerland: Karger.

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition, 137*, 196–209. doi: 10.1016/j.cognition.2015.01.013

Koop, G. J. (2013). An assessment of the temporal dynamics of moral decisions. *Judgment and Decision Making, 8*, 527–539.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica: Journal of the Econometric Society, 28*, 287–309. doi: 10.2307/1907722

Körner, A., & Volk, S. (2014). Concrete and abstract ways to deontology: Cognitive capacity moderates construal level effects on moral judgments. *Journal of Experimental Social Psychology, 55*, 139–145. doi: 10.1016/j.jesp.2014.07.002

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition, 108*, 754–770. doi: 10.1016/j.cognition.2008.06.009

Laham, S. M., Alter, A. L., & Goodwin, G. P. (2009). Easy on the mind, easy on the wrongdoer: Discrepantly fluent violations are deemed less morally wrong. *Cognition, 112*, 462–466. doi: 10.1016/j.cognition.2009.06.001

Lieberman, M. D. (2013). *Social: Why our brains are wired to connect*. New York, NY: Crown.

Lombrozo, T. (2009). The role of moral commitments in moral judgment. *Cognitive Science, 33*, 273–286. doi: 10.1111/j.1551-6709.2009.01013.x

Ludwiczack, A. (2016). *The value of effort: How do we make decisions when effort is involved?* (Unpublished thesis). London, UK: Queen Mary University of London.

Martel, M. M., Nigg, J. T., Wong, M. M., Fitzgerald, H. E., Jester, J. M., Puttler, L. I., ... Zucker, R. A. (2007). Childhood and adolescent resiliency, regulation, and executive functioning in relation to adolescent problems and competence in a high-risk sample. *Development and Psychopathology, 19*, 541–563. doi: 10.1017/S0954579407070265

McGuire, J., Langdon, R., Coltheart, M., & Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology, 45*, 577–580. doi: 10.1016/j.jesp.2009.01.002

McKinnon, M. C., & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition, 102*, 179–218. doi: 10.1016/j.cognition.2005.12.011

Mikhail, J. (2011). *Elements of moral cognition. Rawls' linguistic analogy and the cognitive science of moral and legal judgement*. New York, NY: Cambridge University Press.

Miller, P. A., Eisenberg, N., Fabes, R. A., & Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Developmental Psychology, 32*, 210–219. doi: 10.1037/0012-1649.32.2.210

Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science, 19*, 549–557. doi: 10.1111/j.1467-9280.2008.02122.x

Morris, S. A., & McDonald, R. A. (1995). The role of moral intensity in moral judgments: An empirical investigation. *Journal of Business Ethics, 14*, 715–726. doi: 10.1007/BF00872325

Neumann, J. V., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition, 84*, 221–236. doi: 10.1016/S0010-0277(02)00048-3

Osman, M. (2015a). Moral judgments: Truth, order and Consequence. *Psychology, 6*, 633–642. doi: 10.4236/psych.2015.65061

Osman, M. (2015b). Dynamic moral judgments and emotions. *Psychology, 6*, 922–931. doi: 10.4236/psych.2015.68090

Paharia, N., Kassam, K. S., Greene, J. D., & Bazerman, M. H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes, 109*, 134–141. doi: 10.1016/j.obhdp.2009.03.002

Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology, 17*, 145–171. doi: 10.1016/0162-3095(96)00041-6

Piaget, J. (1965). *The moral judgment of the child* (original 1932). New York, NY: Free Press.

Powell, N. L., Derbyshire, S. W., & Guttentag, R. E. (2012). Biases in children's and adults' moral judgments. *Journal of Experimental Child Psychology, 113*, 186–193. doi: 10.1016/j.jecp.2012.03.006

Pratt, M. W., Diessner, R., Pratt, A., Hunsberger, B., & Pancer, S. M. (1996). Moral and social reasoning and perspective taking in later life: A longitudinal study. *Psychology and Aging, 11*, 66–73. doi: 10.1037/0882-7974.11.1.66

Prelec, D., & Loewenstein, G. (1991). Decision-making over time and under uncertainty – A common approach. *Management Science, 37*, 770–786. doi: 10.1287/mnsc.37.7.770

Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *The Behavioral and Brain Sciences, 4*, 371–388. doi: 10.1017/S0140525X00009407

Rai, T. S., & Holyoak, K. J. (2010). Moral principles or consumer preferences? Alternative framings of the trolley problem. *Cognitive Science, 34*, 311–321. doi: 10.1111/j.1551-6709.2009.01088.x

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews: Neuroscience, 9*, 545–556. doi: 10.1038/nrn2357

Regan, C. (2014). Decisions, dopamine, and degeneracy in complex biological systems. *Neuroscience and Neuroeconomics, 11*, 11–18. doi: 10.2147/nan.s32234

Rest, J. R. (1983). Morality. In P. H. Mussen (Ed.), *Handbook of child psychology: Vol. 3. Cognitive development* (4th ed., pp. 556–629). New York, NY: Wiley.

Rigoux, L., & Guigon, E. (2012). A model of reward- and effort-based optimal decision making and motor control. *PLoS Computational Biology, 8*, e1002716. doi: 10.1371/journal.pcbi.1002716

Ritov, I., & Baron, J. (1992). Status-quo and omission biases. *Journal of Risk and Uncertainty, 5*, 46–61. doi: 10.1007/BF00208786

Ritov, I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes, 79*, 79–94. doi: 10.1006/obhd.1999.2839

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1*, 7–59.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition, 124*, 325–333. doi: 10.1016/j.cognition.2012.06.004

Schmidt, M. F., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science, 21*, 232–236. doi: 10.1016/j.cognition.2012.06.004

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*, 1096–1109. doi: 10.1177/0146167208317771

Shaw, T. (2016). The psychologists take power. *The New York Review of Books*. Retrieved from http://www.nybooks.com/articles/2016/02/25/the-psychologists-take-power/

Shenhav, A., & 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

Singhapakdi, A., Vitell, S. J., & Kraft, K. L. (1996). Moral intensity and ethical decision-making of marketing professionals. *Journal of Business Research, 36*, 245–255. doi: 10.1016/0148-2963(95)00155-7

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology, 27*, 76–105. doi: 10.1016/0022-1031(91)90011-T

Sripada, C. S., & Stich, S. (2005). A framework for the psychology of norms. In P. Carruthers (Ed.), *The innate mind: Culture and cognition* (pp. 280–301). Oxford, UK: Oxford University Press.

Stein, E. W., & Ahmad, N. (2009). Using the analytical hierarchy process (AHP) to construct a measure of the magnitude of consequences component of moral intensity. *Journal of Business Ethics, 89*, 391–407. doi: 0.1007/s10551-008-0006-8

Stewart, L., & Pascual-Leone, J. (1992). Mental capacity constraints and the development of moral reasoning. *Journal of Experimental Child Psychology, 54*, 251–287. doi: 10.1016/0022-0965(92)90020-7

Sugarman, S. D. (1989). *Doing away with personal injury law: New compensation mechanisms for victims, consumers, and business*. New York, NY: Quorum Books.

Tassy, S., Oullier, O., Duclos, Y., Coulon, O., Mancini, J., Deruelle, C., ... Wicker, B. (2011). Disrupting the right prefrontal cortex alters moral judgement. *Social Cognitive and Affective Neuroscience, 7*, 282–288. doi: 10.1093/scan/nsr008

Trémolière, B., De Neys, W., & Bonnefon, J. F. (in press). Reasoning and moral judgment: A common experimental toolbox. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning*. Hove, UK: Psychology Press.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*, 31–48. doi: 10.1037/h0026750

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty, 5*, 297–323. doi: 10.1007/BF00122574

Tykocinski, O. E., Pittman, T. S., & Tuttle, E. E. (1995). Inaction inertia: Foregoing future benefits as a result of an initial failure to act. *Journal of Personality and Social Psychology, 68*, 793–803. doi: 10.1037/0022-3514.68.5.793

Vartanian, O., Mandel, D. R., & Duncan, M. (2011). Money or life: Behavioral and neural context effects on choice under uncertainty. *Journal of Neuroscience, Psychology, and Economics, 4*, 25–36. doi: 10.1037/a0021241

Vera-Estay, E., Dooley, J. J., & Beauchamp, M. H. (2014). Cognitive underpinnings of moral reasoning in adolescence: The contribution of executive functions. *Journal of Moral Education, 44*, 17–33. doi: 10.1080/03057240.2014.986077

Vera-Estay, E., Seni, A. G., Champagne, C., & Beauchamp, M. H. (2016). All for one: Contributions of age, socioeconomic factor, executive functioning, and social cognition to moral reasoning in childhood. *Frontiers in Psychology, 7*, 227. doi: 10.3389/fpsyg.2016.00227

Waldmann, M. R., Nagel, J., & Wiegmann, A. (2012). Moral judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 364–389). Oxford, UK: Oxford Library of Psychology.

Walker, L. (2004). Gus in the gap: Bridging the judgment-action gap in moral functioning. In D. Lapsley & D. Narvaez (Eds.), *Moral development, self and identity* (pp. 1–20). Mahwah, NJ: LEA.

Weber, J. (1996). Influences upon managerial moral decision making: Nature of the harm and magnitude of consequences. *Human Relations, 49*, 1–22. doi: 10.1177/001872679604900101

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*, 780–784. doi: 10.1111/j.1467-9280.2005.01614.x

Wiegmann, A., Horvath, J., & Meyer, K. (under review). Intuitive expertise and irrelevant options.

Wiegmann, A., Okan, Y., & Nagel, J. (2012). Order effects in moral judgment. *Philosophical Psychology, 25*, 813–836. doi: 10.1080/09515089.2011.631995

Wiegmann, A., & Waldmann, M. R. (2014). Transfer effects between moral dilemmas: A causal model theory. *Cognition, 131*, 28–43. doi: 10.1016/j.cognition.2013.12.004

Magda Osman
Queen Mary University of London
Biology and Experimental Psychology Centre
Mile End Road
London, E14NS
UK
m.osman@qmul.ac.uk