Rev.Phil.Psych. (2012) 3:405–423
DOI 10.1007/s13164-012-0112-5

Two Minds Vs. Two Philosophies: Mind Perception Defines Morality and Dissolves the Debate Between Deontology and Utilitarianism

Kurt Gray · Chelsea Schein

Published online: 18 September 2012
© Springer Science+Business Media B.V. 2012

Abstract Mind perception is the essence of moral judgment. Broadly, moral standing is linked to perceptions of mind, with moral responsibility tied to perceived agency, and moral rights tied to perceived experience. More specifically, moral judgments are based on a fundamental template of two perceived minds—an intentional agent and a suffering patient. This dyadic template grows out of the universal power of harm, and serves as a cognitive working model through which even atypical moral events are understood. Thus, all instances of immorality are perceived to involve both blameworthy agents (i.e., acts) and suffering victims (i.e., consequences). Because moral cognition simultaneously concerns acts and consequences, theories which focus primarily on acts (i.e., deontology) or consequences (i.e., utilitarianism) do not accurately describe moral cognition. Indeed, the phenomenon of dyadic completion suggests that deontological and utilitarian concerns are not only simultaneously active, but also typically compatible and reinforcing: wrong acts have harmful consequences, and harmful consequences stem from wrong acts. The cognitive fusion of acts with consequences suggests that normative conflicts between deontology and utilitarianism are not reflected in everyday moral judgment. This in turn suggests that empirical conclusions drawn from moral dilemmas that pit utilitarianism against deontology—i.e., trolley problems—give an skewed account of moral cognition.

In a far away kingdom, a prince rides through a thick forest. Unlike neighboring countries, with their long list of laws, this kingdom has only one rule, and subjects are free to do what they wish as long as they do not break it. The same rule also applies to the prince, but being royalty has its advantages, and few would dare to challenge him if he were caught violating it. Nevertheless, when he breaks the rule—as he often does—he does so as discreetly as possible.

K. Gray (✉) · C. Schein
Department of Psychology, University of North Carolina—Chapel Hill, CB #3270 Davie Hall, Chapel Hill, NC 27599, USA
e-mail: kurtgray@unc.edu

Dusk is slowly falling in the forest, when a flash of color catches his eye. He slows his horse and sees a young girl with blue eyes and a ragged yellow dress hiding amongst the trees. He orders her to come out and asks her name.

"Viola. I'm walking to find work in the next town."

"But the next town is miles away," the prince replies. "Where are your parents?"

"They passed away. I'm an orphan."

The prince smiles slyly, "So you're alone." After looking around to make sure they are truly alone, the prince decides that tonight, he will again be breaking the kingdom's one rule.

One thing is clear in this story—the prince violated an important norm—and it seems likely that the girl is somehow involved. As for the exact nature of this norm violation, the possibilities are limitless: the kingdom may prohibit horseback hitch-hiking, camping in the woods or mingling outside your social class. For all the facts of the story tell us, the prince could be clandestinely adopting this orphan, rescuing her from a life of drudgery so she may grow up joyously in his castle. Nevertheless, we cannot help feeling that Viola will be harmed instead of helped, that she will suffer physically and emotionally at the hands of the rule-breaking prince.

The question is: Why do we tie this norm violation to suffering? In this paper, we suggest that judgments of immorality are generally—and inextricably—tied to perceived suffering. This psychological link between *blame* and *pain* grows out of the fundamental cognitive template unifying moral judgment—the moral dyad (Gray et al. 2012b, c). The moral dyad consists of two perceived minds—an intentional moral agent and a suffering moral patient—and it explains not only a host of phenomena in moral psychology, but also dissolves the psychological tension between deontology and utilitarianism.

At its heart, *dyadic morality* is a formalization of the general link between mind perception and moral judgment, which we explore first in this paper. We then explore dyadic morality, explaining how it grows naturally out of the structure of psychological concepts. Next, we explore the implications of dyadic morality, including dyadic completion, which is the tendency to see harm in immorality, and immorality in harm. Finally, we argue that dyadic morality psychologically unites the theories of utilitarianism and deontology. People do not separate acts from consequences (as deontology advocates), nor do they separate consequences from the agents who perform them (as utilitarianism advocates). This link between deontology and utilitarianism suggests that research in moral cognition goes off track when focusing primarily on moral dilemmas that pit the two theories against each other.

1 Minds and Moral Judgments

Philosophers have long wrestled with two basic questions of morality: who can be held morally accountable and who deserves moral rights. In answering these questions, philosophers often appeal to mental capacities. Moral responsibility is argued to

require the mental capacities of higher-order decision-making (Aristotle, 1111b5-1113b3), second order volitions (Frankfurt 1971), or strong evaluations (Taylor 1985). Moral rights have also been argued to hinge on mental capacities. Kant (1997) suggested that, just as rationality is a prerequisite for moral responsibility, so too is it necessary for moral rights. Bentham (1879; see also Singer 1975), instead emphasized the importance of the capacity for pain, arguing "the question is not, Can they reason? nor, Can they talk? but, Can they suffer?"

Despite disagreement on the exact mental requirements for moral rights or moral responsibility, there is generally consensus among philosophers and legal scholars that mind matters—entities with more sophisticated minds (e.g., humans) are generally argued to deserve greater moral standing than entities with less sophisticated minds (e.g., insects). Empirical accounts also support the link between mind and morality, but they challenge the psychological utility of a number of assumptions made by philosophical accounts.

The first assumption challenged is that there is an objective fact about whether an entity possesses a mind. Scholars have long argued about whether machines can think (Searle 1980; Turing 1950) whether animals can feel pain (DeGrazia 1996; Harrison 1991) and whether vegetative patients possess personality (Singer 1996), but the inaccessibility of other minds makes such questions ultimately unanswerable. Instead, questions of mind perception hinge upon perception (Epley and Waytz 2009), a subjective feeling of "someone being home" when we look into their eyes. Although such perceptions generally track objective tests of mental abilities, they can diverge widely—for example, people will view vegetative patients as having less mind than someone who is dead (Gray et al. 2011b). Despite the fact that mind perception is labile and inconsistent, depending on situational (Epley et al. 2007) and individual differences (Waytz et al. 2010), research reveals it to be the basis for moral judgments (Gray et al. 2012c). Different people may perceive different amounts of mind in a fetus, but those who see highly developed fetal minds invariably also ascribe it important moral rights (Gray et al. 2007). Of course, the powerful link between mind and morality suggests that moral judgments are as labile and variable as mind perception—and research reveals wide variation across cultures in moral judgment (Haidt et al. 1993); accordingly, there is wide cultural variation in mind perception (Cohen and Gunz 2002; Guthrie 1995).

1.1 Mind Perception (and Moral Standing) is Two Dimensional

The second philosophical assumption that empirical research challenges is that mind progresses along a single continuum from no mind (e.g., rocks) to full mind (e.g., humans)—ordering entities along a great chain of being (Brandt and Reyna 2011; Nee 2005). Instead, our research (Gray et al. 2007, 2011a; Gray and Wegner 2012) finds that minds are perceived along the two independent dimensions of agency (the capacity to do, act and plan) and experience (the capacity to feel and sense). This two dimensional structure is advocated by a number of researchers. Robins and Jack (2006) suggest that perceivers can take an intentional stance or a phenomenal stance concerning entities, and Knobe and Prinz (2008) distinguish between the ascription of non-phenomenal states and phenomenal states. Huebner (2009) makes a similar distinction, as do Sytsma and Machery (2009) when they separate non-affective

states (neutral) from affective states (good/bad).[1] A similar two dimensional structure is also found in perceptions of humanness (Haslam 2006; Haslam et al. 2008) and stereotypes (Fiske et al. 2007), and in the distinction between cognitive and emotional empathy (Shamay-Tsoory et al. 2009).

In a somewhat different take, Arico et al. (2011) argue that low-level agency (e.g. eye movement, self-propelled motion) predisposes people to ascribe all manner of mental states to entities. This predisposition is compatible with a two-dimensional model of mind perception, however, as once an entity is identified as capable of mental states, the ascription of agency and experience can progress independently. It is also important to note that low-level agency is quite different from the high-level agency discussed so far (Gray et al. 2007), and we suggest that the qualities focused on by Arico et al. (2011) might better be called *animacy*. Indeed, the term "animated beings" is used by Heider and Simmel (1944) in their experiment investigating the anthropomorphization of self-propelled shapes.

Departing from a dimensional model, Buckwalter and Phelan (unpublished) suggest that mind perception is driven primarily by analytic functionalist concerns, whereby entities are ascribed mental states consistent with their behavior. This perspective is also consistent with two dimensions, as different classes of functions may give rise to different kinds of mind perception. Indeed one researcher (Hallgren, this issue), suggests that two different needs of perceivers—dictating two different functions of perceived entities—can lead to perceptions of either experience or agency. In particular, a need for social connection increases perceived experience, whereas a need for control increases perceived agency (Hallgren, this issue, Epley et al. 2007).

Despite the different ways to conceptualize mind perception, a key benefit of the two-dimensional structure of agency and experience is that it predicts judgments of moral rights and moral responsibilities. In contrast to suggestions that ratings of rights and responsibility are nested one within the other in a single moral circle (see chapter 5 in Bloom 2004 for an overview of these positions),[2] data reveals that these two senses of moral standing are independent—just like perceptions of mind (Gray et al. 2007). This means that although some entities may be ascribed both rights and responsibility (e.g., adult humans), others are ascribed only moral rights (e.g., puppies), or only moral responsibility (e.g., God). Importantly, judgments of moral responsibility are tied mostly to perceptions of agency (r=0.82) and not experience (r=0.22), whereas judgments of moral rights are tied mostly to perceptions of experience (r=0.85) and not agency (r=0.26; but see also Sytsma & Machery, this issue). In the parlance of philosophy, experience qualifies an entity as a moral patient, whereas agency qualifies an entity as a moral agent.[3]

Further support for two distinct links between moral responsibility and agency, and moral rights and patiency comes from those with disordered mind perception (Gray et al. 2011a). Those higher on the autism-spectrum perceive reduced agency (Baron-Cohen

[1] As affect is grounded by pleasure and displeasure/pain (Russell 1980)—core capacities of experience—affective states are analogous to experience; conversely, agentic states are inherently non-affective.

[2] Classically, it was thought that all those with moral responsibility possess moral rights, but not all those with moral rights possess responsibility. As an analogy, all doctors can be patients, though not all patients have the requisite knowledge to be doctors.

[3] Moral agency can also be split by valence into blame and praise (Bastian et al. 2011), harm and help (Gray and Wegner 2011a), and proscriptive and prescriptive (Janoff-Bulman et al. 2009)

et al. 1985) and accordingly give abnormal blame judgments (Moran et al. 2011). Complementarily, those higher on psychopathy have impaired emotional empathy and emotion recognition (Blair 2005; Mahmut et al. 2008; Marsh and Blair 2008) and accordingly fail to see entities such as babies or animals as worthy of moral rights (Glenn et al. 2009; Koenigs et al. 2007).

In general, then, mind perception and moral judgment appear to be tightly tied together. Where there is perceived agency, there is also moral agency (and vice versa); and where there is perceived experience, there is also moral patiency (and vice versa). In the next section, we will see that mind is not only tied to morality, but also defines the very core of moral judgment.

2 The Moral Dyad: Two Minds Psychologically Define Moral Judgment

How would you define sports (or games, Wittgenstein 2001)? As decades of research in philosophy and psychology has found, an exact and all-encompassing definition of anything is hard to find (Murphy 2004). Nonetheless it is still possible to come up with a fuzzy or probabilistic definition that represents the most essential features of something. A definition of sports might be "formalized competition between two or more people (or groups of people), in which one person (or group) is the winner and the rest are losers." This definition may miss the nuances of individual sports—neglecting the finesse of fencing, the camaraderie of curling, and the brutality of boxing—but it nonetheless achieves an abstract and parsimonious understanding.

We suggest that morality can also be characterized by a similar definition involving two parties, and that immoral deeds are characterized by a "perceived intentional agent causing perceived suffering to a moral patient." In other words, morality can be defined by the dyadic interaction of two perceived minds—a moral agent and a moral patient (Gray et al. 2012c). This may not be the dictionary definition of immorality, but we suggest that it is the *psychological* definition—that the mind possesses a template for understanding morality based upon a dyad of agent and patient.

Indeed, all concepts are psychologically understood, not as lists of logical conditions, but instead as fuzzy templates (prototypes or exemplar sets) that abstractly represent the most frequent, distinguishing or pragmatically important elements of a category (for a review, see Murphy 2004). For example, the psychological template of "bird" is something small, feathered, flying and seed-eating because these characteristics best overarch individual examples and best distinguish them from other creatures. See Fig. 1. Importantly, the presence of any specific counter-examples—whether flight-less penguins, large ostriches, and carrion eating-vultures—does not invalidate the template because psychological templates are fuzzy, with no hard and fast lines of inclusion or exclusion (Murphy 2004). Instead, examples vary continuously on the extent to which they fit the template (Rosch and Mervis 1975), with more prototypical examples (e.g., robins) categorized more quickly and free-recalled more frequently than non-prototypical examples (e.g., vultures).

For morality, we suggest that the importance (van Leeuwen and Park 2011), cross-cultural prevalence (Haidt 2007) and affective power of intentional harm (Stotland 1969)—i.e., acts involving an intentional agent and suffering patient—leads the

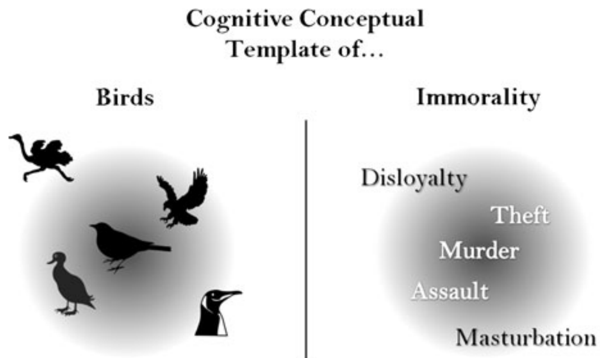

Fig. 1 Concepts are represented not as lists or definition but as fuzzy templates. More prototypical examples are closer to the center; they are better recalled and help to shape subsequent conceptual inferences

human mind to construct a dyadic template of morality. The presence of a template suggests that dyadic acts should be most prototypical of immorality, just as robins are most prototypical of birds. Indeed, when participants—whether American or Indian—are asked to list an immoral act, they overwhelming list a dyadic act (e.g., murder) rather than a non-dyadic act (e.g., masturbation; Gray et al. 2012a).

Importantly, conceptual templates are not only abstract representations of the most canonical and important elements of categories, but also serve as cognitive working models—lenses though which all examples are viewed (Craik 1967). See Fig. 2. These cognitive working models exert top-down influence when categorizing new examples, whether they are canonical or not (DeScioli et al. 2012; Ditto et al. 2012; Gray et al. 2012b). In the case of birds, if you read about a newly discovered bird with scales and teeth, the sparrow-like template of bird would influence your cognition and make you likely to picture the bird as nonetheless scale-less and toothless. We suggest that a dyadic moral template also serves as a cognitive working model, turning even non-canonical moral events into dyadic interactions (Fig. 2). This means that ostensibly victimless moral transgressions (e.g., flag burning) are still perceived to harm victims. In other words, when people are presented with an agent who violates moral norms (e.g., the Prince from the beginning of the paper), the dyadic templates compels them to see suffering moral patients (e.g., poor Viola). We call this phenomena *dyadic completion.*[4]

2.1 Dyadic Completion

Take a look at the evil, venom-dripping snake in Fig. 3. Notice anything else in the picture? You likely cannot help but see the snake encircling a crying baby. Such visual completion occurs because of the top-down influence of higher-level gestalt

[4] Dyadic completion is one implication of a dyadic template. Another is moral typecasting, which is the tendency to see others as *either* moral agents *or* moral patients (Gray and Wegner 2009). Typecasting is one reason why victims escape blame (Gray and Wegner 2011b), why people are willing to harm the saintly (Gray and Wegner 2009), and why good and evil deeds make people physically stronger (Gray 2010). For a review, see Gray et al. (2012c).

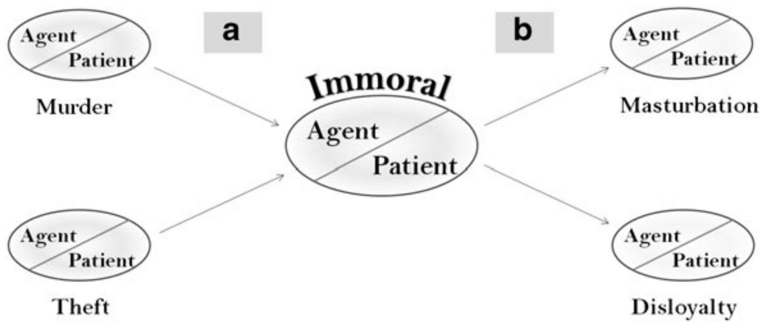

Fig. 2 A dyadic moral template grows out of the power of harm-based moral exemplars (**a**). This template then serves as a cognitive working model for viewing all moral violations (**b**), meaning that even apparently non-dyadic acts appear dyadic

principles, and we suggest the existence of similar top-down influences in moral cognition. In particular, a gestalt representation of morality as dyadic leads people to see both an agent and a patient in immoral deeds and when one is lacking, they are compelled to psychologically complete the dyad. Thus, when there is unjust suffering without an apparent agent, people nonetheless perceive an intentional agent (Gray and Wegner 2010; Mandelbaum and Ripley, this issue); and when there is an immoral agent without an apparent victim, people nonetheless perceive a suffering patient. For example, dyadic completion suggests that if you judge a homosexual lifestyle to be immoral, you perceive it to have a victim (Bryant 1977). Of course, one could debate the "objective" presence of such victims, but such debates are beside the point. Just as with judgments of mind—and those of immorality—the presence of victims is a matter of perception,[5] and we simply suggest that perceptions of victimhood are bound to those of immorality.

Evidence for dyadic completion comes from a number of studies. In one paradigm, participants are asked whether a variety of ostensibly harmless transgressions (e.g., burning a flag) were a) immoral and b) harmed a victim. As predicted, the more acts are rated as immoral, the more they are seen as involving a victim (DeScioli et al. 2012; Gray et al. 2012a). Importantly, dyadic completion occurs not only with white, Western, liberals (WEIRD populations; Henrich et al. 2010), but also with Indians, Brazilians and conservative Americans (Gray et al. 2012a; Haidt et al. 1993; Shweder et al. 1997). One potential question about dyadic completion is whether it is simply a product of post-hoc justification—seeing suffering to rationalize moral intuitions (Haidt 2001). Evidence argues against this idea, as experiments reveal that ostensibly harmless concerns are implicitly linked to harm in cases when justification is unlikely or impossible (Gray et al. 2012a).

Broadly, the fact that various moral transgressions remain psychologically linked to harm undercuts the idea that moral judgment is divided into functionally distinct modules (Haidt and Joseph 2004; Sinnott-Armstrong and Wheatley 2011). Although there are clearly descriptively different moral concerns that account for important individual and cultural differences in judgment (Ditto and Koleva 2011; Graham et al.

[5] The perceptual nature of victimhood casts doubt on evidence that uses "victimless" scenarios (e.g., Haidt 2001). Just because researchers do not see victims does not mean participants do not.

Fig. 3 A visual representation of dyadic completion (see also the Kanizsa triangle). Just as the visual system cannot help but see the baby encircled by the snake, a dyadic moral template means that people cannot help but see moral patients in response to moral agents

2009), these concerns appear to be fundamentally unified by a dyadic template.[6] Indeed, we suggest that various concerns (e.g., dishonesty, disloyalty) represent the specific *content* of the interaction between agent and patient—the precise way in which the agent is perceived to cause suffering to the patient—whereas the dyad represents the fundamental cognitive *form* of morality (see Gray et al. 2012b).

More narrowly, dyadic completion suggests that people cannot psychologically separate immoral acts from the suffering of victims.[7] It is the reason why we cannot shake the feeling that poor Viola suffers at the hands of the prince in the opening story. As we will soon see, it is also the reason why dyadic morality dissolves the psychological debate between deontology and utilitarianism.

3 Dyadic Morality Dissolves the Psychological Debate Between Deontology and Utilitarianism

Traditionally, the line between moral psychology and moral philosophy was clear. Psychology examined *is* questions, developing descriptive theories about how people make moral decisions. In contrast, philosophy examined *ought* questions, developing normative theories about how people *should* make moral decisions. Recently, this sharp line between *is* and *ought* has been blurred as normative theories are advanced to describe how people make moral judgments. In particular, researchers have argued that the two theories of deontology and utilitarianism represent two distinct and competing modes of moral decision-making (e.g., Greene et al. 2001). Evidence for this idea comes from (so-called) trolley problems—moral dilemmas where an unpalatable *act* (e.g., murder) is pitted against desirable *consequences* (e.g., saving lives). Because deontology emphasizes acts over consequences, and utilitarianism emphasizes consequences over acts, these two normative theories aptly describe the two conflicting intuitions in these moral dilemmas. The question, however, is whether

[6] Additionally arguing for unification, different cultures (i.e., liberals and conservative) are also both equally likely to possess moral conviction (Skitka and Bauman 2008)

[7] Indeed, dyadic completion occurs even for the experience of pain and pleasure (Gray 2012).

these competing normative theories describe moral cognition beyond the bounds of trolley problems.

We suggest that just as one can commit the naturalistic fallacy by moving from *is* to *ought*, researchers may have committed the reverse fallacy by moving too quickly from *ought* to *is*, mistaking philosophical theories for psychological reality. In particular, we suggest that the longstanding normative conflict between utilitarianism and deontology traps researchers into a distorted view of moral cognition—a view which presumes an enduring psychological conflict between acts and consequences. In this section we suggest that—within typical moral cognition—deontological concerns and utilitarian concerns are neither independent from each other nor are they competing. Instead, dyadic morality—and in particular dyadic completion—suggests that concerns about acts and consequences are bound together and mutually reinforcing. A dyadic moral template means that concerns about acts (i.e., the deeds of moral agents) are inextricably bound to those about consequences (i.e., the suffering of moral patients). In concrete terms, this means that people generally see bad acts as causing bad consequences, and bad consequences as stemming from bad acts (see Liu and Ditto 2012). This psychological fusion of acts and consequences means that neither act-focused deontology nor consequence-focused utilitarianism offer a complete account of moral cognition. Moreover, it suggests that paradigms which force the conflict of these two normative theories—i.e., trolley problems—give a skewed account of moral cognition.

In this section, we first review how neither deontology nor utilitarianism alone provides a full account for moral cognition. We then explore the fundamental psychological connection between deontological- and utilitarian-oriented moral cognition, arguing against the usefulness of trolley problems as a method to investigate morality.

3.1 Deontology

At its core, deontology claims that the right is independent from the good. For the pure deontologist,[8] an act of lying that saves a life is just as impermissible as an act of lying that takes a life, because it is the act itself that is moral or immoral, regardless of consequences. In the framework of dyadic morality, deontology concerns the moral agent's acts and duties—do not lie, do not steal, do not kill—and deems irrelevant the consequences that befall the moral patient.[9] Claiming that judgments of moral agents are separate from the suffering of moral patients, however, neglects the empirical evidence that agent-centered and patient-centered concerns are psychologically inseparable.

As explored in the previous section, a dyadic template serves as a cognitive working model that psychologically binds together agents and patients. Thus, in typical moral misdeeds, such as murder, theft and lying, agent evaluations are inseparable from concerns about the suffering of victims of murder, theft or dishonesty. Indeed, extensive

[8] There are a variety of deontological accounts (see Alexander and Moore 2008 for an overview). Here we focus on what most accounts share in common—the focus on moral rights and duties over consequences.

[9] There are patient-centered deontological accounts, which focus on moral rights such as treating people not as a means to an end, but as ends in and of themselves. These accounts, however, still fail because violations of moral right are not necessarily linked to consequences—i.e., the experience of suffering.

evidence demonstrates that moral intuitions are sensitive to the suffering of moral patients (Cushman et al. 2006; Turiel 1983). Even young children, who seem like perfect deontologists because of their reliance on explicit rules (Kohlberg 1981; Piaget 1932), will judge a moral principle as immoral if it causes harm—even if it is endorsed by God (Nucci and Turiel 1993). On the opposite spectrum, psychopaths will judge typically impermissible acts (i.e., kicking a dog) as allowable (Glenn et al. 2009), not because they lack an understanding of moral principles, but because they are incapable of understanding the suffering of moral patients (Gray et al. 2011a; Marsh and Blair 2008). Where suffering is perceived, so too is moral impermissibility.

If judgments of typical misdeeds are tied to the suffering of victims, what of those concerning atypical misdeeds such as masturbation, flag burning and personal drug use? Because even atypical misdeeds are understood through a harm-based dyadic template, their perceived wrongness remains powerfully tied to perceived suffering. Converging evidence linking consequences to wrongness is provided by Ditto and Liu (2011), who suggest that deontologists rely on a “consequentialist crutch” when making moral judgments. For instance, those who believe in the permissibility of capital punishment also believe that it reduces the murder rate, whereas those who believe in its impermissibility not only see no effect on the murder rate, but also see it as irrevocably harming those innocently convicted (Liu and Ditto 2012). Ditto and Liu (2011) explain these effects as motivated justification of deontic judgments, suggesting perhaps that once the need to justify is eliminated, the link between morality and consequences is severed. Contradicting this possibility, dyadic morality finds that this link is deep-seated, implicit, and automatic, meaning that people simply cannot separate the acts of agents from their influence on moral patients (Gray et al. 2012c).

Thus, in contrast to Ditto and Liu (2011), Haidt (2001) and Greene (2007), we deny that judgments of wrongness are more fundamental than those of harm—instead, our research suggests that wrongness and harm are fundamentally bound together as two sides of the same moral coin. Indeed, if moral judgments were meaningfully distinct from consequences then a discrepancy between moral judgments and perceived outcome should be common. However, it is exceptionally rare for people to believe that an act is both morally wrong and that it leads to good outcomes (Liu and Ditto 2012). Those who oppose condom use, for example, do not believe that it curbs STDs or lowers teen pregnancies, but instead inevitably link its wrongness to physical, emotional and spiritual suffering.

One possible objection about this critique of deontology is that judgments of moral wrongness can persist even when a scenario (e.g., incest) has been explicitly rendered “harmless,” ostensibly suggesting that the suffering of moral patients is psychologically separable from moral judgments (Haidt 2001). First, there are obvious ways in which these harmless scenarios remain linked to objective harm (Gray et al. 2012b; Huebner 2011; Jacobson 2012). Second, they provide character cues (Inbar et al. 2012; Pizarro and Tannenbaum 2011), which in turn predict more harmful (and typical) moral acts. Most importantly, the question is whether explicit denials of suffering succeed in dispelling the *implicit* perceptions of suffering that drive moral judgments. This is doubtful, as people invariably bring up concerns of harm to explain their moral judgment despite explicit experimenter avowals of harmlessness

(Haidt et al. 2000). Furthermore, under cognitive load, concerns about harm become *more* salient, not less (Gray et al. 2012a).

Indeed, extensive research casts doubt on the power of explicit conscious processes to influence moral judgment (Haidt 2001), and so we suggest that explicit denials of harm are useless because implicit perceptions of harm are so deeply rooted in moral cognition (Gray et al. 2012a). To illustrate the distinction between explicit and implicit perceptions of harm, consider the scenario used by Gendler (2008): Imagine you are standing on the Grand Canyon Skywalk, a transparent platform jutting out over the canyon, and when you look down and see only thousands of feet of nothing beneath you. In this case, you may explicitly acknowledge your safety, but as your pounding heart attests, you cannot shake the implicit feeling of harm. Analogously, even though an experimenter tells you a particular case of incest is safe, you cannot buck millions of years of evolutionary pressure telling you otherwise.[10]

In sum, we suggest that people do not psychologically engage in a strict form of deontology because moral violations are inextricably tied to perceived harm. Such perceived harm is not only a consequence of moral judgment, but also a cause, whereby perceived wrongness is sensitive to concerns about suffering—whether explicit, implicit, or potential. Acts committed by moral agents are no doubt important, but they are not psychologically judged in isolation; the *right* is not separated from the *good*.

3.2 Utilitarianism

Deontologists may focus on moral agents at the expense of consequences, but the reverse is not necessary true for utilitarians. Utilitarians may focus primarily on the consequences of moral actions—on the realized or potential suffering of moral patients—but they also acknowledge the importance of agent-based factors leading up to these outcomes. Nevertheless, we suggest this acknowledgement is insufficient, because people care about moral agents even when such concerns are irrelevant—or even antagonistic—to concerns about consequences.

The key characteristic of moral agents is moral agency—the capacity for intention—and perceived intention is clearly a key determinant of moral judgment (Heider 1958; Shaver 1985). The importance of intentions is generally accepted by utilitarians,[11] but usually only as a cause of moral responsibility. Research demonstrates, however, that perceived intention is not only a cause of moral judgment but also a consequence (Knobe 2003), and that the valence of intentions can actually influence the experience of moral acts—malicious intentions make electric shocks hurt more (Gray and

[10] More explicitly, we suggest that the conclusion typically reached from moral dumbfounding (Haidt et al. 2000) is analogous to the following fear-of-heights logic. <<You are terrified of heights. I tell you that you are completely safe. You begrudgingly acknowledge that this is true, but remain terrified. Given this explicit acknowledgment, I surmise that there must be some additional non-height related reason for your fear.>> In truth, there is no additional fear-inducing factor—you are simply unable to intuitively accept the lack of danger.

[11] For example, Sidgwick (1907) writes, “moralists of all schools … would agree that the moral judgments … relate primarily to intentional actions … on this point I do not think that any dispute need arise.”

Wegner 2008) and benevolent intentions make massages feel better (Gray 2012). Importantly, studies demonstrate that people care about intentions themselves, above and beyond their causal connection to outcomes.

In an emerging line of research on character, researchers find that participants will ascribe blame for acts that are relatively less harmful (Pizarro et al. 2012; Tannenbaum et al. 2011) or entirely harmless (Inbar et al. 2012). For example, people will ascribe blame to a stock-broker who stands to profit from a natural disaster—even though he did not directly cause harm—simply because of his wicked desires (Inbar et al. 2012).[12] People also ascribe moral impermissibility to morally neutral choices, as two studies conducted in our lab demonstrate. In the first study, we gave half of the participants a modified footbridge scenario: a train is heading towards five workers, but you can save them by flipping a switch, which in turn drops a man through a trapdoor onto the tracks. In this condition, 46 out of 61 of the participants (75 %) selected flipping the switch as the morally acceptable action.

In the other condition, participants were given a similar scenario, but were told that there were *two* footbridges—a red and a blue one—each with a man on them, and that they could flip one of two switches—either the red or blue one. Even though the choice was consequentially the same as the other condition—kill one or five—significantly fewer participants (33/57 or 58 %) saw flipping *either* switch as morally acceptable, $\chi^2=4.09$, $p<0.04$. We suggest that this choice over who to kill, although irrelevant to the mathematics of suffering, adds perceived agency to the scenario, and therefore decreases its moral acceptability. In other words, because morality is understood through a dyad of intention+suffering, increasing agency increases perceived wrongness when holding objective suffering constant.

In a follow-up, we presented participants with the scenario in Fig. 4, and told them that the switch next to them selected which track the trolley would take on its way to kill 5 people. Unsurprisingly, most participants (12/66 or 82 %) said that they would prefer not to pick the track, but importantly, the majority (41/66 or 62 %) also said that picking the track was immoral,[13] despite the fact that it was consequentially irrelevant to outcomes. People appear to be attuned to the mere presence of intention, even when it cannot affect utility.

One final testimony to the importance of agency is that dyadic completion occurs not only in response to agents, but also in response to patients, whereby people perceive an intentional agent to account for perceived unjust suffering (Gray et al. 2012c; Mandelbaum & Ripley, this volume; Nichols and Knobe 2007). Anecdotally, this can be seen in the impulse of people to sue others after accidents in which they feel unjustly harmed from slipping on ice or tripping on sprinkler heads (Andrews 1981). Experimentally, people will also see agents in response to suffering, and when they cannot blame another person they will often look to animals (Oldridge 2004) or God (Gray and Wegner 2010). It would be one thing if people blame agents after misfortune and leave it there, but historical accounts document extremely effortful

[12] Of course, dyadic completion suggests that harm—whether past, present or potential—remains perceived. Indeed, people will connect intention to harm even when capable only through the magic of voodoo (Pronin et al. 2006)

[13] 41/66 is significantly higher than chance, $\chi^2=3.88$, $p<0.05$. Furthermore, judgments of immorality are significantly different from judgments of whether they would do it, $\chi^2=6.86$, $p<0.05$, suggesting that participants appreciated the difference between these two questions.

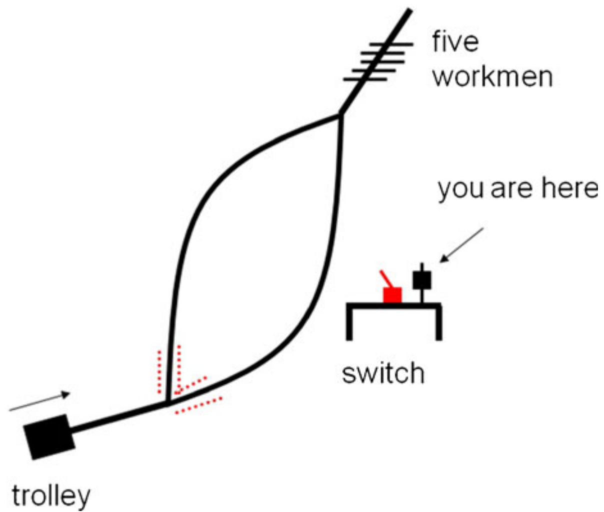

Fig. 4 Consequentially irrelevant choices influence moral judgment

behavior to punish perceived agents. For example, in medieval France, peasants would not only hold livestock responsible for otherwise agentless harms, but would also put them on trial and—if found guilty—would build a custom gallows from which to hang them (Humphrey 2003; Oldridge 2004).

These behaviors argue against the idea that people can make moral decisions according to strict utilitarianism. After a child has been eaten by a pig there is no additional utility in bringing the pig to the magistrate for trial (Oldridge 2004), but the enduring link between suffering and agency suggests that people have a powerful need to blame someone after a moral transgression. Indeed, the sheer effort it takes to prosecute and hang a pig suggests that people can care more about (perceived) intentions than utilitarian concerns—it's not like the hanging will deter future pigs from child-eating. More prosaically, empirical studies also find that people will place intentions before consequences, as they will ascribe punishment based on retribution at the expense of utilitarian concerns such as deterrence (Carlsmith 2006). Furthermore, in both judgment and behavior, people are motivated to act morally above and beyond the consequences of those acts (Bauman et al. 2012; Skitka and Bauman 2008)

Just as with deontology, people do not engage in strict utilitarianism; they care not only about suffering, but are also deeply concerned with moral agents. Research finds that people pass judgment on seemingly inconsequential choices, and will ignore consequential factors in identifying, blaming and punishing moral agents. The moral dyad prompts people to care simultaneously about acts and agents, pain and patients, suggesting that—psychologically—people rely on elements of both deontological and utilitarian accounts.

3.3 An Uneasy Marriage: Deontology and Utilitarianism

John Gottman studies marriage by forcing conflict. He brings couples into the lab and makes them fight over contentious topics, such as in-laws, sex or spending habits.

These antagonistic snap-shots yield some important information about the stability of the relationship (Carrere et al. 2000), but it would be misleading to equate marriages with conflict. Couples may occasionally fight, but relationships are generally better characterized by interdependence, affinity, cooperation and connection—the very opposite of conflict

In moral psychology, researchers also seek to make general claims about moral cognition by observing snap-shots of conflict between deontological and utilitarian ideas. Typical moral dilemmas, for instance, pit concerns about collective welfare (e.g., 5 vs. 1) against those of violating specific norms (e.g., murder), and ask subjects to chose between them. Just like the fights between Gottman's couples, these scenarios can be useful,[14] but whether they can yield truths about moral cognition in general depends on whether they are representative of most moral situations. Specifically, the question is whether deontological and utilitarian concerns conflict—not just in philosophical discussions—but also in everyday psychology. If they do not, then it suggests that moral dilemmas are limited in their power to offer universal insights about moral cognition.

Dyadic morality suggests that, psychologically, deontological and utilitarian intuitions are like a husband and wife: in certain circumstances they will bicker, but they are generally characterized by unity—where one is found, the other will also be found, and they tend to mutually reinforce each other. The moral dyad binds together the acts of moral agents with their consequences upon moral patients, suggesting that people simultaneous care about deontological and utilitarian concerns. Blame for an act cannot be considered in isolation from its harmful consequences; and harmful consequences cannot be considered in isolation from an intentional agent. Turning again to the work of Ditto and Liu (2011), they find that deontic judgments of morality consistently map on to utilitarian judgments. In real life, bad acts are perceived to have bad consequences, and bad consequences are perceived to stem from bad acts; seldom are bad acts perceived to results in good consequences, or are good consequences perceived to come from bad acts. In contrast, moral dilemmas always pit bad (e.g., murder) acts against good consequences (e.g., saving lives), which is a combination seldom perceived to occur in important everyday moral domains such as birth control, capital punishment, racial integration, off-shore drilling or torture (Ditto and Liu 2011). Of course, many real life scenarios do objectively involve deontological and utilitarian trade-offs, but our point is that these trade-offs are rendered *psychologically invisible* by the dyadic template. Objectively, torture may violate individual rights for collectively useful information, but those who oppose torture perceive this information to be useless, and those who advocate torture fail to ascribe moral rights to its victims (Greenberg and Dratel 2005)

One objection to this claim of compatibility is fMRI evidence demonstrating conflicts between cognitive (utilitarian) and affective (deontological) brain areas in moral decision making (Greene et al. 2001). Philosophers (Berker 2009; Kumar and Campbell 2012) and psychologists (McGuire et al. 2009) raise some questions about these conclusions (answered by Greene 2010), but an additional point bears mentioning—namely that this research all occurs within the domain of trolley problems. These moral dilemmas are not only designed to pit utilitarianism against deontology,

[14] indeed, we used two trolley scenarios in the previous section!

but also emotion against cognition, which is typically not the case in moral cognition. Research by Haidt and colleagues, for instance, finds that cognition typically serves to explain emotional responses (Haidt 2001; Wheatley and Haidt 2005), and Hume (1740) suggests that cognition cannot help but serve affect at some level. The cooperation between affect and cognition is further suggested by sophisticated brain connectivity analyses that demonstrate that both are underlain by similar—and often mutually reinforcing—neural networks (Pessoa 2008).[15]

In sum, the fact that utilitarianism and deontology—and affect and cognition—can be made to conflict does not mean that they typically do. Dyadic morality, along with other research, points to the interdependence and cooperation of both utilitarian and deontological intuitions, and affect and cognition. This frequent cooperation means that we should be cautious in making universal claims about moral cognition from dilemmas specifically designed to make them conflict. Trolley problems are a powerful tool in the moral psychologist's toolbox, but they give a skewed view of moral cognition because they ignore the enduring psychological link between acts and consequences, agents and patients.

4 Conclusion

Since the time of Aristotle, it has been clear that mind is linked to moral standing. Less clear, however, is exactly *how* mind is linked to morality. In this paper, we have explored this question in detail, linking two dimensions of mind (agency and experience) to two kinds of moral standing (moral agents and moral patients), and have also emphasized that questions of mind—just like those of morality—are matters of perception.

The distinction of moral agents and patients not only provides a link to mind perception, but also provides the basis for a fundamental moral template. This dyadic template of agent/patient is initially abstracted from canonically immoral acts involving direct harm, and subsequently serves as a general lens through which all morality is understood, even atypical acts such as purity or group-oriented violations. The dyadic template, while simplifying various moral contents into one single and parsimonious form, also unifies apparently distinct modes of moral cognition. In particular, it suggests frequent psychological agreement between deontological- and utilitarian-oriented intuitions, and highlights the importance of moving beyond psychological paradigms that pit them against each other.

Returning to the story at the beginning, we may never know what the prince did to Viola on that summer evening, but research suggests that we automatically bind together perceived immorality and perceived harm. What is more, the evil Prince and vulnerable Viola are not simply characters in a vignette, but represent the two archetypal roles in the cognitive structure of morality—an intentional agent and a suffering patient bound inextricably together.

[15] Pessoa (2008) writes "The current view…[is that] many [brain] regions can be conceptualized as either 'affective' or 'cognitive'. Popular examples are the amygdala in the domain of emotion and the lateral prefrontal cortex in the case of cognition. This prevalent view is problematic for a number of reasons. Here, I will argue that complex cognitive–emotional behaviours have their basis in dynamic coalitions of networks of brain areas, none of which should be conceptualized as specifically affective or cognitive."

References

Alexander, L., and M. Moore. 2008. Deontological ethics. In *The Stanford Encyclopedia of Philosophy,* ed. E.N. Zalta (Fall 2008.). Retrieved from http://plato.stanford.edu/archives/fall2008/entries/ethics-deontological/.

Andrews, L.B. 1981. Suing as a first resort: A review of Marks's The Suing of America and Lieberman's The Litigious Society. *American Bar Foundation Research Journal* 6(3): 851–860.

Arico, A., B. Fiala, R.F. Goldberg, and S. Nichols. 2011. The folk psychology of consciousness. *Mind & Language* 26(3): 327–352. doi:10.1111/j.1468-0017.2011.01420.x.

Baron-Cohen, S., A.M. Leslie, and U. Frith. 1985. Does the autistic child have a "theory of mind"? *Cognition* 21(1): 37–46.

Bastian, B., S.M. Laham, S. Wilson, N. Haslam, and P. Koval. 2011. Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology* 50(3): 469–483. doi:10.1348/014466610X521383.

Bauman, C.W., D.C. Wisneski, and L.J. Skitka. 2012. Cubist consequentialism: The pros and cons of an agent–patient template for morality. *Psychological Inquiry* 23(2): 129–133. doi:10.1080/1047840X.2012.668014.

Bentham, J. 1879. *An introduction to the principles of morals and legislation*. Oxford: The Clarendon Press.

Berker, S. 2009. The normative insignificance of neuroscience, the normative insignificance of neuroscience. *Philosophy & Public Affairs* 37(4): 293–329. doi:10.1111/j.1088-4963.2009.01164.x.

Blair, R.J.R. 2005. Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition* 14(4): 698–718. doi:10.1016/j.concog.2005.06.004.

Bloom, P. 2004. *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.

Brandt, M.J., and C. Reyna. 2011. The chain of being: A hierarchy of morality. *Perspectives on Psychological Science* 6(5): 428–446. doi:10.1177/1745691611414587.

Bryant, A. 1977. *The Anita Bryant story: The survival of our nation's families and the threat of militant homosexuality*. Grand Rapids: Revell.

Carlsmith, K.M. 2006. The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology* 42(4): 437–451. doi:16/j.jesp.2005.06.007.

Carrere, S., K.T. Buehlman, J.M. Gottman, J.A. Coan, and L. Ruckstuhl. 2000. Predicting marital stability and divorce in newlywed couples. *Journal of Family Psychology* 14(1): 42.

Cohen, D., and A. Gunz. 2002. As seen by the other…: Perspectives on the self in the memories and emotional perceptions of Easterners and Westerners. *Psychological Science* 13(1): 55–59.

Craik, K.J.W. 1967. *The nature of explanation*. Cambridge: Cambridge University Press.

Cushman, F., L. Young, and M. Hauser. 2006. The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science* 17(12): 1082–1089. doi:10.1111/j.1467-9280.2006.01834.x.

DeGrazia, D. 1996. *Taking Animals Seriously: Mental Life and Moral Status*. Cambridge: Cambridge University Press.

DeScioli, P., S. Gilbert, and R. Kurzban. 2012. Indelible victims and persistent punishers in moral cognition. *Psychological Inquiry* 23(2): 143–149.

Ditto, P.H., and S.P. Koleva. 2011. Moral empathy gaps and the American culture war. *Emotion Review* 3 (3): 331–332. doi:10.1177/1754073911402393.

Ditto, P.H., and B. Liu. 2011. Deontological dissonance and the consequentialist crutch. In *The social psychology of morality: Exploring the causes of good and evil*, ed. M. Mikulincer and P.R. Shaver. Washington: APA Press.

Ditto, P.H., B. Liu, and S.P. Wojcik. 2012. Is anything sacred anymore. *Psychological Inquiry* 23(2): 155–161.

Epley, N., and A. Waytz. 2009. Mind perception. In *The handbook of social psychology*, 5th ed, ed. S.T. Fiske, D.T. Gilbert, and G. Lindzey, 498–541. New York: Wiley.

Epley, N., A. Waytz, and J.T. Cacioppo. 2007. On seeing human: A three-factor theory of anthropomorphism. *Psychological Review* 114(4): 864–886.

Fiske, S.T., A.J.C. Cuddy, and P. Glick. 2007. Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences* 11(2): 77–83.

Frankfurt, H.G. 1971. Freedom of the will and the concept of a person. *The Journal of Philosophy* 68(1): 5–20.

Gendler, T.S. 2008. Alief and belief. *Journal of Philosophy* 105(10): 634–663.

Glenn, A.L., R. Iyer, J. Graham, S. Koleva, and J. Haidt. 2009. Are all types of morality compromised in psychopathy? *Journal of Personality Disorders* 23(4): 384–398. doi:10.1521/pedi.2009.23.4.384.

Graham, J., J. Haidt, and B.A. Nosek. 2009. Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology* 96(5): 1029–1046. doi:10.1037/a0015141.

Gray, K. 2010. Moral transformation: Good and evil turn the weak into the mighty. *Social Psychological and Personality Science* 1(3): 253–258. doi:10.1177/1948550610367686.

Gray, K. 2012. The power of good intentions: Perceived benevolence soothes pain, increases pleasure, and improves taste. *Social Psychological and Personality Science* 3: 639–645.

Gray, K., and D.M. Wegner. 2008. The sting of intentional pain. *Psychological Science* 19(12): 1260–1262. doi:10.1111/j.1467-9280.2008.02208.x.

Gray, K., and D.M. Wegner. 2009. Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology* 96(3): 505–520. doi:10.1037/a0013748.

Gray, K., and D.M. Wegner. 2010. Blaming God for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review* 14(1): 7–16. doi:10.1177/1088868309350299.

Gray, K., and D.M. Wegner. 2011a. Dimensions of moral emotions. *Emotion Review* 3(3): 227–229.

Gray, K., and D.M. Wegner. 2011b. To escape blame, don't be a hero–Be a victim. *Journal of Experimental Social Psychology* 47(2): 516–519. doi:10.1016/j.jesp.2010.12.012.

Gray, K., and D.M. Wegner. 2012. Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition* 125(1): 125–130. doi:10.1016/j.cognition.2012.06.007.

Gray, H.M., K. Gray, and D.M. Wegner. 2007. *Dimensions of mind perception. Science* 315(5812): 619.

Gray, K., A.C. Jenkins, A.S. Heberlein, and D.M. Wegner. 2011a. Distortions of mind perception in psychopathology. *Proceedings of the National Academy of Sciences* 108(2): 477–479. doi:10.1073/pnas.1015493108.

Gray, K., T.A. Knickman, and D.M. Wegner. 2011b. More dead than dead: Perceptions of persons in the persistent vegetative state. *Cognition* 121(2): 275–280. doi:10.1016/j.cognition.2011.06.014.

Gray, K., C. Schein, and A.F. Ward. 2012a. *The harm hypothesis: Perceived harm unifies morality.* Unpublished manuscript, University of North Carolina.

Gray, K., A. Waytz, and L. Young. 2012b. The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry* 23: 206–215.

Gray, K., L. Young, and A. Waytz. 2012c. Mind perception is the essence of morality. *Psychological Inquiry* 23: 101–124.

Greenberg, K.J., and J.L. Dratel (eds.). 2005. *The torture papers: The road to Abu Ghraib*. Cambridge: Cambridge University Press.

Greene, J.D. 2007. The secret joke of Kant's soul. In *Moral psychology*, The neuroscience of morality: Emotion, disease, and development, vol. 3, ed. W. Sinnott-Armstrong, 35–80. Cambridge: MIT Press.

Greene, J.D. 2010. Notes on "The normative insignifigance of neuroscience" by Selim Berker. Retrieved from http://www.wjh.harvard.edu/~jgreene/GreeneWJH/Greene-Notes-on-Berker-Nov10.pdf

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

Guthrie, S.E. 1995. *Faces in the clouds: A new theory of religion*. New York: Oxford University Press.

Haidt, J. 2001. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review* 108(4): 814–834.

Haidt, J. 2007. The new synthesis in moral psychology. *Science* 316(5827): 998–1002. doi:10.1126/science.1137651.

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

Haidt, J., S.H. Koller, and M.G. Dias. 1993. Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology* 65(4): 613–628.

Haidt, J., F. Bjorklund, and S. Murphy. 2000. *Moral dumbfounding: When intuition finds no reason*. Unpublished manuscript, University of Virginia.

Harrison, P. 1991. Do animals feel pain. *Philosophy* 66(255): 25–40.

Haslam, N. 2006. Dehumanization: An integrative review. *Personality and Social Psychology Review* 10(3): 252–264. doi:10.1207/s15327957pspr1003_4.

Haslam, N., S. Loughnan, Y. Kashima, and P. Bain. 2008. Attributing and denying humanness to others. *European Review of Social Psychology* 19: 55. doi:10.1080/10463280801981645.

Heider, Fritz. 1958. *The psychology of interpersonal relations*. New York: Wiley.

Heider, F., and M. Simmel. 1944. An experimental study of apparent behavior. *American Journal of Psychology* 57: 243–259. doi:10.2307/1416950.

Henrich, J., S. Heine, and A. Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2–3): 61–83.
Huebner, B. 2009. Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies? *Phenomenology and the Cognitive Sciences* 9(1): 133–155. doi:10.1007/s11097-009-9126-6.
Huebner, B. 2011. Critiquing empirical moral psychology. *Philosophy of the Social Sciences* 41(1): 50–83. doi:10.1177/0048393110388888.
Hume, D. 1740. *A treatise of human nature*. New York: Dover.
Humphrey, N. 2003. *The mind made flesh: Essays from the frontiers of psychology and evolution*. New York: Oxford University Press.
Inbar, Y., D.A. Pizarro, and F. Cushman. 2012. Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin* 38(1): 52–62.
Jacobson, D. 2012. Moral dumbfounding and moral stupefaction. *Oxford Studies in Normative Ethics*, *2*. (forthcoming).
Janoff-Bulman, R., S. Sheikh, and S. Hepp. 2009. Proscriptive versus prescriptive morality: Two faces of moral regulation. *Journal of Personality and Social Psychology* 96(3): 521–537. doi:10.1037/a0013779.
Kant, I. 1997. *Lectures on ethics*. Cambridge, UK: Cambridge University Press.
Knobe, J. 2003. Intentional action and side effects in ordinary language. *Analysis* 63: 190–193.
Knobe, J., and J. Prinz. 2008. Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences* 7(1): 67–83. doi:10.1007/s11097-007-9066-y.
Koenigs, M., L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, and A. Damasio. 2007. Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446(7138): 908–911. doi:10.1038/nature05631.
Kohlberg, L. 1981. *The philosophy of moral development: Moral stages and the idea of justice*. New York: HarperCollins.
Kumar, V., and R. Campbell. 2012. On the normative significance of experimental moral psychology. *Philosophical Psychology* 1–20. doi:10.1080/09515089.2012.660140.
Liu, B.S., and P.H. Ditto. 2012. What dilemma? *Moral evaluation shapes factual belief. Social Psychological and Personality Science.*. doi:10.1177/1948550612456045.
Mahmut, M.K., J. Homewood, and R.J. Stevenson. 2008. The characteristics of non-criminals with high psychopathy traits: Are they similar to criminal psychopaths? *Journal of Research in Personality* 42 (3): 679–692. doi:10.1016/j.jrp.2007.09.002.
Marsh, A.A., and R.J.R. Blair. 2008. Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews* 32(3): 454–465. doi:10.1016/j.neubiorev.2007.08.003.
McGuire, J., R. Langdon, M. Coltheart, and C. Mackenzie. 2009. A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology* 45(3): 577–580.
Moran, J.M., L.L. Young, R. Saxe, S.M. Lee, D. O'Young, P.L. Mavros, and J.D. Gabrieli. 2011. Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences* 108(7): 2688–2692. doi:10.1073/pnas.1011734108.
Murphy, G.L. 2004. *The big book of concepts*. Cambridge: MIT Press.
Nee, S. 2005. The great chain of being. *Nature* 435(7041): 429–429. doi:10.1038/435429a.
Nichols, S., and J. Knobe. 2007. Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs* 41(4): 663–685. doi:10.1111/j.1468-0068.2007.00666.x.
Nucci, L., and E. Turiel. 1993. God's word, religious rules, and their relation to Christian and Jewish children's concepts of morality. *Child Development* 64(5): 1475–1491. doi:10.1111/j.1467-8624.1993.tb02965.x.
Oldridge, D.J. 2004. *Strange histories: The trial of the pig, the walking dead, and other matters of fact from the medieval and renaissance worlds*. New York: Routledge.
Pessoa, L. 2008. On the relationship between emotion and cognition. *Nature Reviews Neuroscience* 9(2): 148–158. doi:10.1038/nrn2317.
Piaget, J. 1932. *The moral judgment of the child*. New York: Harcourt Brace.
Pizarro, D.A., and D. Tannenbaum. 2011. Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In *The social psychology of morality: Exploring the causes of good and evil*, ed. M. Mikulincer and P.R. Shaver, 91–108. Washington, DC: APA Press.
Pizarro, D.A., D. Tannenbaum, and E. Uhlmann. 2012. Mindless, harmless, and blameworthy. *Psychological Inquiry* 23(2): 185–188. doi:10.1080/1047840X.2012.670100.

Pronin, E., D.M. Wegner, K. McCarthy, and S. Rodriguez. 2006. Everyday magical powers: The role of apparent mental causation in the overestimation of personal influence. *Journal of Personality and Social Psychology* 91(2): 218–231. doi:10.1037/0022-3514.91.2.218.
Robbins, P., and A.I. Jack. 2006. The phenomenal stance. *Philosophical Studies* 127(1): 59–85. doi:10.1007/s11098-005-1730-x.
Rosch, E., and C.B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7(4): 573–605. doi:10.1016/0010-0285(75)90024-9.
Russell, J.A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39(6): 1161–1178. doi:10.1037/h0077714.
Searle, J.R. 1980. Minds, Brains and Programs. Reprinted in *The Philosophy of Artificial Intelligence*, ed. M. Boden, 67–88. Oxford: Oxford University Press, 1990.
Shamay-Tsoory, S.G., J. Aharon-Peretz, and D. Perry. 2009. Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132(3): 617–627.
Shaver, K.G. 1985. *The attribution of blame: Causality, reponsibility, and blameworthiness*. New York: Springer.
Shweder, R.A., N.C. Much, M. Mahapatra, and L. Park. 1997. The “big three” of morality (autonomy, community, and divinity), and the “big three” explanations of suffering. In *Morality and Health*, eds. A. Brandt and P. Rozin, 119–169. New York: Routledge.
Sidgwick, H. 1907. *The methods of ethics*. Indianapolis: Hackett.
Singer, P. 1975. *Animal liberation*. New York: Random House.
Singer, P. 1996. *Rethinking Life and Death: The Collapse of Our Traditional Ethics* (Reprint.). New York: St. Martin’s Griffin.
Sinnott-Armstrong, W., and T. Wheatley. 2011. *Moral judgments are not unified: Why moral psychology needs a taxonomic approach*. Manuscript under review, Duke University.
Skitka, L.J., and C.W. Bauman. 2008. Moral conviction and political engagement. *Political Psychology* 29 (1): 29–54. doi:10.1111/j.1467-9221.2007.00611.x.
Stotland, E. 1969. Exploratory investigations of empathy. In *Advances in experimental social psychology*, vol. 4, ed. L. Berkowitz, 271–314. New York: Academic.
Sytsma, J., and E. Machery. 2009. Two conceptions of subjective experience. *Philosophical Studies* 151(2): 299–327. doi:10.1007/s11098-009-9439-x.
Tannenbaum, D., E.L. Uhlmann, and D. Diermeier. 2011. Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology* 47: 1249–1254.
Taylor, C. 1985. *Human agency and language*. Cambridge: Cambridge University Press.
Turiel, E. 1983. *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.
Turing, A.M. 1950. Computing machinery and intelligence. *Mind, New Series* 59(236): 433–460.
van Leeuwen, F., and J.H. Park. 2011. *Moral virtues and political orientation: Assessing priorities under constraint*. Unpublished Manuscript, University of Bristol.
Waytz, A., J.T. Cacioppo, and N. Epley. 2010. Who sees human? *Perspectives on Psychological Science* 5 (3): 219–232. doi:10.1177/1745691610369336.
Wheatley, T., and J. Haidt. 2005. Hypnotic disgust makes moral judgments more severe. *Psychological Science* 16(10): 780–784. doi:10.1111/j.1467-9280.2005.01614.x.
Wittgenstein, L. 2001. *Philosophical investigations: The German text, with a revised English translation*. (G. E. M. Anscombe, Trans.). Oxford; Malden, Mass.: Blackwell.

Author’s Note

For helpful comments we thank Brendan Ritchie. For Fig. 3, we thank Aaron Scott.

Reproduced with permission of copyright owner.
Further reproduction prohibited without permission.