Psychology Press
Taylor & Francis Group

TARGET ARTICLE

Mind Perception Is the Essence of Morality

Kurt Gray
Department of Psychology, University of North Carolina, Chapel Hill, North Carolina

Liane Young
Department of Psychology, Boston College, Chestnut Hill, Massachusetts

Adam Waytz
Kellogg School of Management, Northwestern University, Evanston, Illinois

Mind perception entails ascribing mental capacities to other entities, whereas moral judgment entails labeling entities as good or bad or actions as right or wrong. We suggest that mind perception is the essence of moral judgment. In particular, we suggest that moral judgment is rooted in a cognitive template of two perceived minds—a moral dyad of an intentional agent and a suffering moral patient. Diverse lines of research support dyadic morality. First, perceptions of mind are linked to moral judgments: dimensions of mind perception (agency and experience) map onto moral types (agents and patients), and deficits of mind perception correspond to difficulties with moral judgment. Second, not only are moral judgments sensitive to perceived agency and experience, but all moral transgressions are fundamentally understood as agency plus experienced suffering—that is, interpersonal harm—even ostensibly harmless acts such as purity violations. Third, dyadic morality uniquely accounts for the phenomena of dyadic completion (seeing agents in response to patients, and vice versa), and moral typecasting (characterizing others as either moral agents or moral patients). Discussion also explores how mind perception can unify morality across explanatory levels, how a dyadic template of morality may be developmentally acquired, and future directions.

In 1945, Pablo Picasso distilled the essence of a bull. In a series of 15 drawings, he took the full complexity of the animal and reduced it to a dozen lines (Figure 1). Although this final sketch does not capture the idiosyncrasies of every particular animal, it remains a compelling representation—a template—of the broader concept. If we attempted to distill the essence of morality, what would result? The full set of "immoral acts" is undeniably complex, and includes murder, cheating, theft, incest, disobedience, and disrespect, to name only a few. Despite this diversity, we suggest that the human mind acts as Picasso did, abstracting out the key elements from various moral transgressions to create a cognitive template. These key elements are *intention* and *pain* (i.e., intentional harm) and the essence of moral judgment is the perception of two complementary minds—a dyad of an intentional *moral agent* and a suffering *moral patient*.

At first blush, dyadic morality may seem too spare to account for the range of moral diversity, but we suggest that—like Picasso's sketch—this dyadic structure represents an important psychological template. The moral dyad can explain, for instance, why some psychological disorders correspond to deficits in both moral judgment and mind perception, why heroic agents are perceived as better able to withstand pain, and even why people believe in God. More broadly, dyadic morality provides a way to unify distinct domains of morality.

Unifying Versus Dividing

For a long time, moral psychology focused on the moral judgment of the child (Kohlberg, 1981; Piaget, 1932; Turiel, 1983) and defined morality as concerns about justice (but see Gilligan, 1993). More recently,
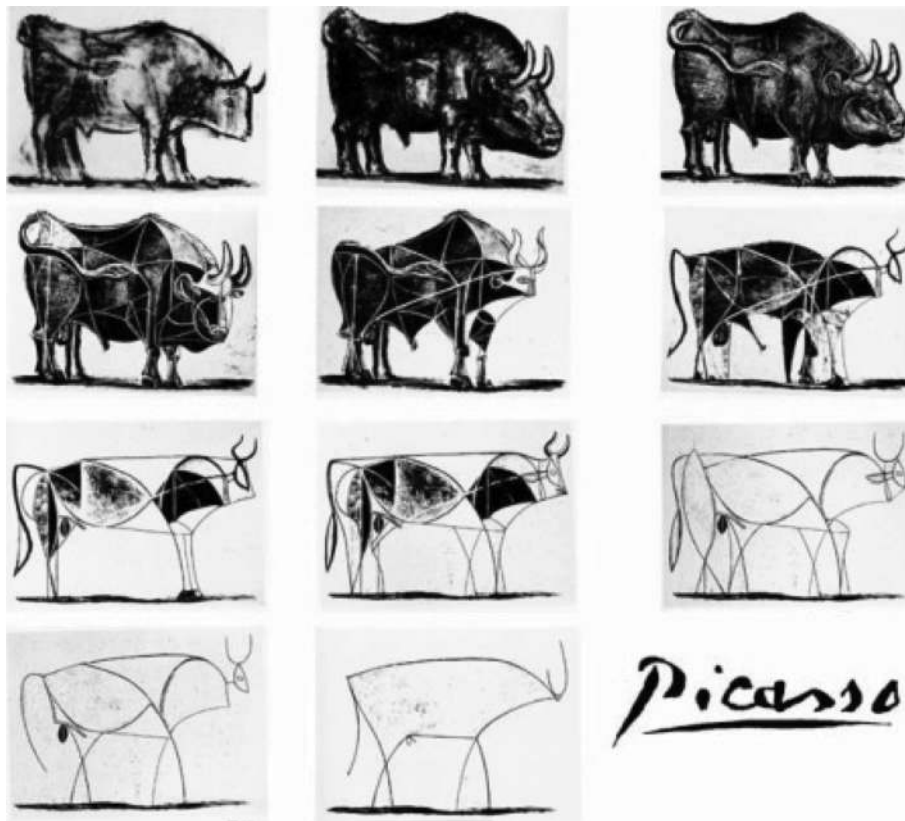
Figure 1. Picasso distills the essence of a bull. © 2012 Estate of Pablo Picasso/Artists Rights Society (ARS), New York, NY.

anthropological research suggests that moral judgment extends beyond justice to include concerns for one's group, one's relationships, and even one's God (Haidt & Graham, 2007; Rai & Fiske, 2011; Shweder, Mahapatra, & Miller, 1987). These findings have led many to focus on the differences between cultures and individuals, rather than common concerns or psychological mechanisms. Some researchers go as far as to suggest that no psychological processes are shared across domains (Parkinson et al., 2011; e.g., Sinnott-Armstrong & Wheatley, 2011).

Questions of similarities and differences, however, often depend on the level of analysis. For example, at a descriptive level, ice and steam are very different, but both are manifestations of the same substance. Likewise, different emotions are associated with different phenomenological experiences as well as expressions and behaviors, but research suggests that all emotions may be explained by appeal to two basic dimensions: valence and arousal (Bliss-Moreau & Barrett, 2009; Russell, 1980). In the case of moral judgment, even descriptively different domains may be unified by an underlying psychological essence. Just as Picasso's bull represents the essence of bulls in general (despite impressive bovine diversity), we advance that morality is essentially represented by a cognitive template that combines a perceived intentional agent with a per-

ceived suffering patient. Before we discuss evidence for this claim, we first define two key terms: *cognitive template* and *perceived*.

Cognitive Templates

Turning from bulls to dogs, we ask, What is the definition of "dog"? This question might seem easy, but a definition that accommodates all dogs is hard to come by. A first guess might be "a barking animal with some fur and a tail," but tailless pugs, hairless Chihuahuas, and voiceless basenjis are all exceptions. In fact, it is difficult to strictly define *anything*, whether it be "dogs," "furniture," or "professors," because concepts are fuzzy and lack defining features (Medin, Wattenmaker, & Hampson, 1987; Murphy, 2004; Rosch & Mervis, 1975; see also Carey, 2009). Nevertheless, the human mind creates its own cognitive templates for concepts, built from the features of exemplars. So despite canine variety, "dog" still conjures to mind an image of a four-legged creature that *does* have fur, that *does* have a tail, and that *does* bark. Similarly, we suggest that, despite the variety of moral transgressions, there is a cognitive template of morality—the moral dyad—which not only integrates across various moral transgressions but also serves as a working model for

understanding the moral world (Baldwin, 1992; Craik, 1967; Johnson-Laird, 1986). This dyadic template fits the majority of moral situations because mind perception is as flexible as moral judgment itself.

Mind Perception

Who or what has a mind? It may be clear to you that you have a mind, but what about other people? Intuitively, it seems obvious that other people have minds too: A friend tells you he is upset, a colleague develops a new theory, a partner plans for the future. Appearances can be deceiving, though—how can you *really* know? Those around you might be "zombies" (Chalmers, 1997), people who, in the philosophical sense, are otherwise indistinguishable from us but lack mental states. Zombies aside, many other entities are also ambiguous—how are we to know whether a fetus, an individual in a persistent vegetative state, or Watson the computer has a mind? Even self-report and measures of brain activity leave other minds ultimately inaccessible and ambiguous. As a consequence, the existence of other minds is a matter of *perception* (Arico, Fiala, Goldberg, & Nichols, 2011; Epley & Waytz, 2010; Gray, Gray, & Wegner, 2007; Huebner, Bruno, & Sarkissian, 2009; Knobe & Prinz, 2008; Malle, 2005). How different people perceive the mind of a single entity can therefore vary tremendously and even defy objective biological evidence of mental capacities (Gray, Knickman, & Wegner, 2011). Thus, for mind perception to be the essence of morality, there need not "objectively" be an intentional agent and suffering patient in every moral situation, but only the perception of this dyad.

In this article, we distinguish mind perception—ascribing a mind to others—from reasoning about the specific contents of those minds. Considerable research has explored the capacity for understanding other minds, targeting theory of mind (Baron-Cohen, Leslie, & Frith, 1985; Gopnik & Wellman, 1992; Premack & Woodruff, 1978; Wimmer & Perner, 1983; Woodward, 1998), mentalizing (Frith & Frith, 2003), and perspective taking (Epley, Caruso, & Bazerman, 2006; Galinsky & Moskowitz, 2000; Stotland, 1969). Although similar cognitive and neural processes may be involved in both perceiving mind and understanding specific mental contents, we suggest that mind perception is more fundamental. Before understanding exactly what someone is thinking or feeling, we must perceive a kind of mind (Epley & Waytz, 2009).

Overview

Many researchers have shown that mental state attribution is important to morality, but here we explore whether mind perception is the *essence* of morality.

We investigate whether all moral judgments can be explained by appealing to a dyadic template: two perceived minds—a moral agent characterized by agency and a moral patient characterized by experience. These perceived minds include people (individuals and groups), animals, robots, and supernatural entities (Epley & Waytz, 2009; Gray et al., 2007). In particular, we investigate whether moral violations are understood primarily in terms of intention and suffering.[1]

First, we describe the dyadic link between mind perception and morality. Then, we explore how distinct moral domains can be understood through the lens of intention and suffering. Next, we explore how the moral dyad compels and constrains moral judgments and perceptions of mind, before discussing how mind perception provides a unifying account for morality across multiple levels of analysis. Finally, we explain how a dyadic template of morality can be acquired and then offer future directions.

Links Between Mind Perception and Morality

Analogous Structures of Mind Perception and Morality

The law has long linked mind perception and morality. For example, those perceived to have reduced mental capacity (e.g., insane persons) are deemed less responsible for their transgressions, and the rights afforded to others hinge on the kind of mind they are ascribed. Empirically, a recent large-scale survey investigated specific links between mind perception and morality. Respondents evaluated both the mental capacities of diverse targets (e.g., adult humans, babies, animals, God) and their moral standing (Gray et al., 2007). In particular, participants assessed whether target entities deserved moral rights and whether they possessed moral responsibility.

Previous proposals suggest that mind perception exists along a single dimension, from inert and mindless (e.g., a rock) to fully functioning and conscious (e.g., a human adult), with both rights and responsibilities increasing along that continuum (Dennett, 1997). Instead, the mind survey revealed that people perceive minds along two independent dimensions. The first dimension, *experience*, is the perceived capacity for sensation and feelings (e.g., hunger, fear, pain, pleasure, and consciousness). The second, *agency*, is the perceived capacity to intend and to act (e.g., self-control, judgment, communication, thought, and memory). An entity can be high on both dimensions (e.g., adult humans), low on experience and high on agency (e.g., God, Google), high on experience and low on agency (e.g., children, animals), or low on both (e.g., the

---

[1]Although moral judgments also apply to good deeds, the bulk of psychology concerns judgments of wrongness or evil (Baumeister, 1999; Zimbardo, 2008), and so we focus primarily on immorality.

deceased, inanimate objects). Other work on mind perception has revealed similar dimensions (Knobe & Prinz, 2008; Kozak, Marsh, & Wegner, 2006; Robbins & Jack, 2006; Sytsma & Machery, 2009), as have cognitive frameworks for stereotype content (Fiske, Cuddy, & Glick, 2007), humanness (Haslam, 2006; Haslam, Loughnan, Kashima, & Bain, 2008), personality (Wiggins & Broughton, 1991), and empathy (Davis, 1996; Decety, 2011).

The mind survey revealed critical links between the dimensions of mind perception and the attribution of moral rights and moral responsibilities (see also Bastian, Laham, Wilson, Haslam, & Koval, in press). Ascriptions of rights were correlated with perceptions of experience, whereas ascriptions of responsibility were correlated with perceptions of agency. In the parlance of philosophy (Aristotle, 2009), agency qualifies entities as moral agents, those who are capable of doing good or evil, whereas experience qualifies entities as moral patients, those who are capable of benefiting from good or suffering from evil (Figure 2). Adult humans usually possess both agency and patiency, and can therefore be both blamed for evil and suffer from it. A puppy, by contrast, is a mere moral patient; we seek to protect it from harm but do not blame it for injustice. Corporations (high in agency, low in patiency) possess the opposite profile (Gray, Gray, & Wegner, 2008; Knobe & Prinz, 2008), possessing responsibility but few rights.

The link between experience and moral rights can explain why abortion and animal experimentation emphasize the experience of such entities; consciousness and emotion confer moral patiency and the right to life (Singer, 1975). In turn, the link between agency and moral responsibility can explain why those who advocate trying adolescents in adult court emphasize their agency; self-control and the capacity to plan allow for the assignment of blame and punishment.

The mind survey demonstrates key connections between mind perception and morality. However, if mind perception is the essence of moral judgment, then deficits in morality and deficits in mind perception should go hand in hand. In the next two sections, we focus on two disorders, one characterized by deficits in mind perception (autism) and the other characterized by deficits in moral behavior (psychopathy). Those who suffer from autism should show corresponding difficulties with moral judgment, whereas those with psychopathy should show corresponding difficulties with mind perception. In addition, we examine whether immoral behavior is linked to reduced mind perception in neurotypical participants.

Autism

Autism is a developmental disorder characterized by difficulty with social interaction (e.g., poor eye contact) and social cognition. Researchers have suggested that the root of autism spectrum disorders, including milder forms (Asperger's Syndrome), is the inability to understand others, or "mindblindness"—the inability to see the minds of others (Baron-Cohen, 1995; Baron-Cohen et al., 1985; Carruthers, 1996). Studies have documented deficits in theory of mind, suggesting that autistic individuals have difficulty inferring others' beliefs and intentions (Happé, 1995; Zalla, Machery, & Leboyer, 2008). Theory of mind difficulties, however, may ultimately stem from deficits of mind perception. If we cannot first perceive the minds of others, then we cannot represent the contents of those minds.

A recent study investigated the link between the autism spectrum and mind perception (Gray, Jenkins, Heberlein, & Wegner, 2011). Participants completed a series of personality measures and the mind survey (Gray et al., 2007). Among the personality measures was the Autism Quotient (AQ), a self-report measure of autism spectrum disorder suitable for assessing the general public (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). As predicted, increased AQ scores were linked to decreased attributions of agency to adult humans (Gray, Jenkins, et al., 2011). Perceptions of mind in other targets were otherwise unchanged, and there was no link between autism and perceptions of experience, which may account for why emotional empathy often remains intact despite difficulty with cognitive perspective taking (Blair, 2005; Smith, 2009).

This inability to attribute agency provides a clear test for the link between mind perception and morality: Individuals with autism should show abnormal patterns of moral responsibility judgments. Indeed, high-functioning adults diagnosed with Asperger's Syndrome deliver aberrant judgments of moral responsibility, assigning abnormally high levels of blame for accidental harms (Moran et al., 2011). The link between mind perception and morality is further demonstrated by an experiment that used transcranial magnetic stimulation (TMS) to interfere neural activity in a region associated with mind perception (right temporo-parietal junction [RTPJ]; Saxe, 2006). Relative to controls, participants who received TMS to the RTPJ judged moral agents abnormally, neglecting their intentions and focusing more on the consequences of their actions (Young, Camprodon, Hauser, Pascual-Leone, & Saxe, 2010).

In parallel, developmental research indicates a link between understanding other minds and moral judgment in neurotypical children. Piaget (1932) first found that young children reason egocentrically (ignoring other people's mental states), and ascribe blame based primarily on outcomes. More recently, research has linked the ability to understand others' false beliefs with ascribing less blame for accidental transgressions

Two dimensions of mind	**Agency**	**Experience**
consisting of...	Intention, Planning	Sensation, Emotion
and also including...	Capacity for Blame	Capacity for Pain
which qualifies entities as...	**Moral Agents**	**Moral Patients**
and confers...	Moral Responsibility	Moral Rights
as illustrated by the disorders of...	Autism	Psychopathy

Figure 2. The correspondence between the two dimensions of mind and the two moral types.

(Killen, Lynn Mulvey, Richardson, Jampol, & Woodward, 2011). In general, once children are able to take mental states into account, they make adultlike moral judgments that account for agents' intentions (Baird & Astington, 2004; Baird & Moses, 2001; Darley & Zanna, 1982; Fincham & Jaspars, 1979; Karniol, 1978; Lane, Wellman, & Evans, 2010; Yuill, 1984). In autism and in typical development, deficits in mind perception correspond to deficits in moral judgment.

Psychopathy

Whether in real life or in the movies, the behavior of a psychopath can be unnerving. Psychopaths are typically callous, manipulative, and indifferent to the suffering of others and commit many of the world's most horrific crimes (Hare, 1998). Psychopaths can inflict harm to achieve their goals but also harm others for pure sport. Childhood narratives of future psychopathic killers, for example, often document the killing or torture of animals (Davis, 1991).

Psychopaths undoubtedly show a distorted sense of morality, but are these moral distortions tied to deficits in mind perception? If psychopaths fail to ascribe moral rights or patiency to others, they should fail to perceive experience in others. Consistent with this idea, psychopaths have difficulty with both emotional empathy and emotional recognition (Blair, 2005; Mahmut, Homewood, & Stevenson, 2008; Marsh & Blair, 2008). Further evidence that psychopaths fail to represent the experience of others is provided by the same large-scale survey previously described (Gray, Jenkins, et al., 2011). In addition to the AQ, participants also completed the Self-Report Psychopathy Scale (Paulhus, Hemphill, & Hare, 2009), which includes subscales of callous affect and manipulation. Psychopaths showed deficits in perceiving experience in others: Higher psychopathy scores were associated with decreased experience ascriptions to adults, children, and animals (Gray, Jenkins, et al., 2011), all entities that psychopaths are more willing to harm (Bartels & Pizarro, 2011; Glenn, Iyer, Graham, Koleva, & Haidt, 2009; Hare, 1998).

The link between ascriptions of experience and moral patiency is further demonstrated by difficulties in moral judgment in those who acquire deficits in experience perception, namely, patients with focal damage to brain regions for social-emotional processing, (e.g., ventromedial prefrontal cortex [VMPFC]). While retaining broader intellectual functioning, these patients (e.g., Phineas Gage) exhibit "acquired sociopathy," with blunted affect and diminished emotional empathy (Anderson, Barrash, Bechara, & Tranel, 2006; Barrash, Tranel, & Anderson, 2000). VMPFC patients are more likely to view inflicting harm in the context of moral dilemmas as more morally acceptable (Ciaramelli, Muccioli, Làdavas, & di Pellegrino, 2007; Glenn et al., 2009; Koenigs et al., 2007), and to view failed attempts to harm as more morally acceptable (Young et al., 2010). Deficits in the perception of experience appear to reduce ascriptions of moral rights and concerns about harming others.

These findings from autism and psychopathy research suggest that mind perception is tied to morality: Deficits in perceiving agency are tied to difficulties in understanding moral agents, whereas deficits in perceiving experience are tied to difficulties in understanding moral patients. This critical link is further demonstrated by the phenomenon of dehumanization.

Moral Disengagement and Dehumanization

We usually like to think of ourselves as virtuous agents, so when we invariably slip up, we must convince ourselves that we're not so bad (Chance, Norton, Gino, & Ariely, 2011; Shu, Gino, & Bazerman, 2011). Cheating may seem wrong, but we can tell ourselves that everyone does it. Likewise, if we harm someone, we can rationalize our actions afterwards by stripping away the victim's mind, because actions are only harmful—and immoral—if someone suffers. In one demonstration of this effect, participants instructed to eat beef jerky later ascribed less mind to cows than those who were instructed to eat cashews (Loughnan, Haslam, & Bastian, 2010).

People also engage in dehumanization—denying mental states to other—to justify acts of aggression or discrimination (Bandura, Barbaranelli, Caprara, &

105

Pastorelli, 1996; Cikara, Eberhardt, & Fiske, 2010; Harris & Fiske, 2006; Haslam et al., 2008; Leyens et al., 2000). In one study, prison guards, inmate support staff members, and executioners reported their attitudes toward inmates. Executioners—those directly involved in the killing of inmates—exhibited the highest levels of dehumanization, suggesting they may have justified their role in capital punishment by denying that their "victims" were moral patients (Osofsky, Bandura, & Zimbardo, 2005). Likewise, when people are reminded of genocide and discrimination against minority groups, they ascribe them less mind in order to reduce the wrongness of these acts and associated guilt (Castano & Giner-Sorolla, 2006; Esses, Veenvliet, Hodson, & Mihic, 2008; Goff, Eberhardt, Williams, & Jackson, 2008). That people dehumanize their victims supports the link between mind perception and morality.

Morality = Agent + Patient

Moral Judgment Should Be Linked to Perceptions of Intention and Suffering

If you've ever played tennis by yourself, hitting a ball against a wall, again and again, you know the feeling that something is missing. In fact, without a partner (or partners), it's hard to even call it tennis. The same goes for morality, which we suggest involves a template of perceived intentional moral agent and a suffering moral patient. If the essence of morality is captured by the combination of harmful intent and painful experience, then acts committed by agents with greater intent and that result in more suffering should be judged as more immoral.

The law assigns more blame for intentional than accidental acts (e.g., murder vs. manslaughter; American Law Institute, 1962), and folk intuitions correspond to legal distinctions: Intentional transgressions are assigned more blame than accidental transgressions (Cushman, 2008; Malle, Guglielmo, & Monroe, in press; Shaver, 1985; Weiner, 1995). Intentions are so powerfully linked to blame that even irrelevant intentions can increase judgments of blame. For example, people forced to kill others at gunpoint are perceived as more immoral when they wanted the man dead, even though they had no choice (Woolfolk, Doris, & Darley, 2006). Unrelated bad intention can also make an act blameworthy, as Alicke (1992) found that people assign more blame for running a stop sign when the driver is rushing home to hide cocaine rather than an anniversary present.

The law also assigns more blame for acts that cause more suffering (e.g., vehicular manslaughter vs. reckless driving; American Law Institute, 1962), and empirical studies also find that blame is linked to the suffering experienced by victims (Cushman, 2008;

Kahneman, Schkade, & Sunstein, 1998; Walster, 1966).[2] Differences in the salience of a suffering victim can also explain the perceived wrongness of crimes such as rape (clear victim) versus tax evasion (unclear victim; Nichols & Knobe, 2007), and also why it is worse not to help identifiable victims (Small & Loewenstein, 2003).

Of course, a dyadic definition of morality suggests that blame is linked to the *combination* of intention and suffering, which suggests that actions with a clear causal link between agent and patient should result in more blame. Indeed, introducing additional causal links in a moral chain diffuses blame (Fincham & Roberts, 1985). In one study with a real-life analogue, people judged a drug company to be less blameworthy when it increased the price of an important drug through an intermediate company than when it did so directly (Paharia, Kassam, Greene, & Bazerman, 2009). People see harm as more permissible when it is inflicted indirectly (Cushman, Young, & Hauser, 2006; Greene, 2007).

The research covered in this section suggests that the combination of intention and suffering increases judgments of moral wrongness, consistent with a dyadic account of morality. Next, we suggest that a dyadic template unifies moral acts across domains.

The Moral Dyad as a General Cognitive Template

The idea that the essence of morality is the perceived interaction between minds echoes other research that emphasizes the social function of morality (Haidt, 2007; Rai & Fiske, 2011). However, defining morality as the combination of intention and suffering may appear to exclude other categories of moral concerns. Anthropology suggests that morality encompasses more than interpersonal harm (Shweder et al., 1987), an idea extended by Haidt, Graham, and colleagues (Graham et al., 2011; Haidt & Graham, 2007) in their model of five independent moral domains: harm, fairness, in-group, authority, and purity. Whereas harm and fairness are directly linked to suffering (Ridley, 1998), concerns for in-group, authority, and purity seem to be independent, revolving around group functioning (Graham & Haidt, 2010). Rai and Fiske (2011) also suggested a broader conception of morality in which moral judgments are determined not by the nature of the act but by the four relationship types of unity, equality, hierarchy, and proportionality. In a similar spirit, Sinnott-Armstrong and Wheatley (2011) denied that harm or any other concept unifies morality.

---

[2]Because humans can easily entertain counterfactuals (Roese, 1997), attempted harm also fits a dyadic template (e.g., attempted murder); the more likely an act is to cause harm, the more immoral it should seem.
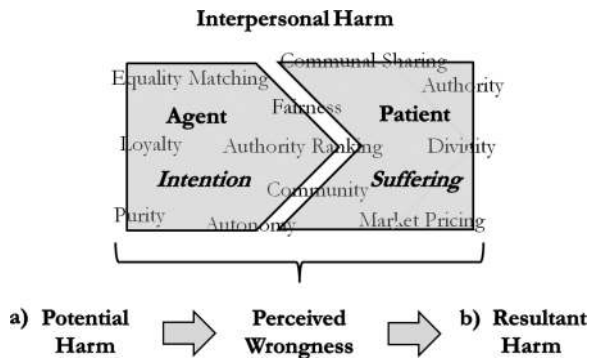
Figure 3. Various moral domains can be understood through the dyadic template of perceived moral agent (intention) and perceived moral patient (suffering), that is, interpersonal harm. *Note.* A link to harm is further demonstrated in two ways: (a) harm related concerns (e.g., perceived danger) increase perceived wrongness and (b) even ostensibly harmless moral violations are linked to resultant harm.

Although these moral taxonomies advocate the presence of a moral agent (one who commits the violation), they do not necessarily recognize the presence of a suffering moral patient. A dyadic template of morality suggests, however, that even these apparently victimless moral acts still involve the perceived presence of a moral patient. This does not mean, of course, that every moral act causes direct physical harm in actuality, but instead that immoral acts lead observers to *perceive* a suffering victim. This suffering can be interpreted through the lens of bodily injury, emotional damage, or even spiritual destruction (Suhler & Churchland, 2011). Indeed, Shweder originally outlined how violations of autonomy, community, or divinity all elicit perceptions of suffering (Shweder, Much, Mahapatra, & Park, 1997). On our account, perceived suffering is not a distinct moral domain, but a core feature of all immoral acts (Figure 3).

A dyadic model of morality makes a number of specific predictions that we develop next concerning the link between various moral domains and perceived suffering. First, not only should it be possible to understand all moral acts in terms of harm and suffering, but general concerns about harm should increase the perceived immorality of acts across all moral domains. Second, people should perceive moral violations across domains as causing suffering. Third, typical moral acts should reflect a dyadic structure. Finally, people should be more concerned with immoral acts that cause direct suffering than those that do not.

Concerns About Suffering Underlie Different Domains

In the old fable of the blind men and the elephant, each man describes a unique experience—whether it be a sinuous trunk or papery ear—but each is actually touching the same animal. We explore how violations of different moral domains each imply harm and suffering, focusing primarily on Haidt's five domains (Haidt, 2007).[3] Instances of harm (e.g., kicking a dog in the head) involve clear suffering, and violations of fairness (e.g., refusing to reciprocate a favor) can cause suffering through depriving others of needed resources. Violations of in-group loyalty (e.g., betrayal) not only cause emotional harm to the betrayed individual but also can lead to physical harm from rival groups who compete against each other for resources. Violations of authority (e.g., disobeying leaders) can also result in suffering. In both human and nonhuman groups, authority structures provide a way of peacefully resolving conflict; violence results when social structures are threatened (Benson & Kugler, 1998; de Waal, 2006; Gould, 2003). Disobeying authority can also be lethal in joint actions, such as when a solider disobeys orders on the battlefield, leaving comrades in danger. Finally, violations of purity—whether related to food or sex—can also lead to suffering. Promiscuous sex can lead to sexually transmitted infections, incest can lead to children with genetic defects, and rancid meat can lead to illness. Impure actions can also result in spiritual suffering, tainting the soul, and offending the gods (Graham & Haidt, 2010; Norenzayan & Shariff, 2008), which can subsequently lead to increased perceptions of physical suffering, (e.g., eternal damnation, more difficult future lives; Kolenda, 1964).

We have outlined how different moral domains may ultimately be rooted in actual (physical, emotional, or spiritual) suffering, but the key question is whether they are *psychologically* linked to suffering. If they are, then general concerns about suffering or harm (i.e., perceived danger) should then increase judgments of moral wrongness. Consistent with this idea, conservatives (relative to liberals) see the world as more dangerous (Jost et al., 2007) and also view many actions as more immoral (Graham, Haidt, & Nosek, 2009). In addition, as people age, they are both more likely to perceive the world as more dangerous (Eibach, Libby, & Gilovich, 2003; Tulloch, 2000) and to moralize concerns about authority and purity (Graham, 2011). Finally, priming death—the ultimate form of harm—increases condemnation of those who violate norms related to in-group, authority, and purity (Greenberg et al., 1990; Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989). In sum, concerns about suffering are linked to judgments of moral wrongness. The next test of dyadic morality is whether moral judgments entail perceptions of suffering.

---

[3]These can be systematically linked to other taxonomies. Specifically, Pinker (2011) linked purity to divinity, loyalty and authority to community, and harm and fairness to autonomy in Shweder's model (Shweder et al., 1987) and links purity and loyalty to communal sharing, authority to authority ranking, harm to equality matching, and fairness to market pricing in Rai and Fiske's model.
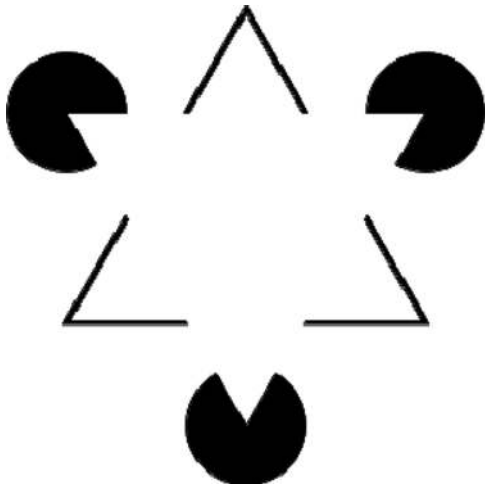
Figure 4. The Kanizsa triangle. *Note.* Just as our minds automatically perceive a downward pointing triangle when presented with this visual template, it is suggested that our minds automatically fill in a dyadic moral template and perceive suffering in contexts of immorality.

Moral Transgressions Involve Perceptions of Suffering

Consider the Kanizsa triangle (Figure 4) three Pac-Man-like objects turn to face each other, and our mind automatically perceives the missing shape. If our cognitive template of morality is dyadic, then the presence of a wrongdoer (a moral agent) should prompt people to perceive a suffering victim, just as our minds fill in the Kanizsa triangle. In other words, even in ostensibly victimless acts, people should complete the moral dyad and perceive a victim. This phenomenon of *dyadic completion* is explored in more detail later, but here we examine the idea that judgments of immorality entail the perception of suffering or harm.

Anecdotal evidence for the link between immorality and perceived suffering is everywhere. For example, Bryant (1977), an antigay activist, wrote a book entitled *The Anita Bryant Story: The Survival of Our Nation's Families and the Threat of Militant Homosexuality*. In it, she suggested that homosexuality not only tears apart families but also irrevocably harms children. Similar sentiments can be found in different cultures; Shweder and colleagues (Shweder et al., 1997) summarized an example from Mumford (1989): "Tibetan communities, for example, have the idea that the malicious or envious gossip of one's neighbors [i.e., blameworthy intention] . . . acts as a kind of force capable of wreaking havoc with one's life and health" (p. 199). Empirical studies by Turiel, Hildebrandt, Wainryb, and Saltzstein (1991) showed that young adults who judged homosexuality and pornography as wrong also perceived these behaviors to causing suffering. Extending this effect, DeScioli (2008) found that individuals who judge harmless deeds (e.g., recreational drug use) as immoral also perceive them to harm victims. Simi-

larly, Royzman, Leeman, and Baron (2009) found that moral judgments of disgusting but harmless acts were linked to perceived harm.

In one study that tested whether moral violations in general involved perceived victims, participants were asked to rate the wrongness of moral transgressions across five moral domains (Graham et al., 2009) and to identify whether a victim was harmed. Not surprisingly, harm violations elicited perceptions of perceived victims, but so did violations of fairness and the ostensibly victimless group-oriented transgressions of in-group, authority, and purity (K. Gray & Ward, 2011). Strikingly, even conservatives saw victims behind these group-oriented violations, despite reports that conservatives possess moral concerns unrelated to harm (Graham et al., 2009). Although perceptions of harm could represent post hoc motivated reasoning (Ditto, Pizarro, & Tannenbaum, 2009; Haidt, 2001; Kunda, 1990; Skitka & Mullen, 2002), all participants were told not to justify their responses, a manipulation that past research has shown eliminates feelings of accountability (Tetlock & Kim, 1987).

A second study examined whether people implicitly tie harm to wrongdoing (Gray & Ward, 2011). Participants read a description of someone burning a flag (an in-group violation) before rating either the painfulness of two injuries (i.e., cutting your finger, stubbing your toe) or the grossness of two foods (i.e., a glass of vinegar, a tin of sardines). The more people judged flag burning as immoral, the more they perceived the injuries as causing suffering (Gray & Ward, 2011). It is important to note that judgments of flag burning were not linked to the perceived grossness of food, arguing against global negative affect driving this effect. Instead, it appears that immorality specifically compels people to see suffering in response to blameworthy intention. Although flag burning may be seen as a symbolic or metaphorical harm, people actually linked this act to physical suffering.

A dyadic template suggests not only that perceived suffering is tied to immorality, but that all morality is understood through the lens of harm. If this is the case, then other moral domains (e.g., purity) should potentiate the concept of harm more than harm should potentiate other moral domains. This was tested by examining whether people are faster to respond to the word *harmful* after being primed with *unfair* (fairness), *disloyal* (ingroup), and *impure* (purity) than vice versa. In other words, if harm is the superordinate concept uniting all of morality, then even ostensibly harmless individual- (fairness) and group-oriented domains (ingroup and purity) should asymmetrically prime harm. Results revealed the predicted pattern, and further found that harm was not potentiated by nonmoral concepts, and that other moral domains did not potentiate each other (Figure 5). This provides additional evidence that moral violations across domains
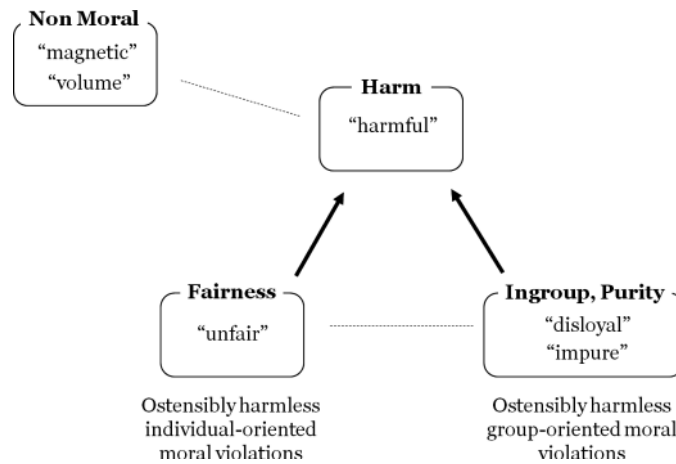
Figure 5. Cognitive associations between various moral domains, as given by asymmetric priming effects. *Note.* Ostensibly harmless domains activate the concept of harm, more than vice versa. The concept of harm was not activated by nonmoral concepts, and non-harm moral domains did not activate each other.

are understood—both explicitly and implicitly—with a dyadic template of blameworthy moral agents and suffering moral patients.

Moral Acts Involving Suffering Are Most Prototypical and Most Important

Moral judgments appear to be tied to a dyadic cognitive template characterized by harm; however, there is no disputing that some acts involve suffering more directly (e.g., murder) than others (e.g., talking back to your parents). A dyadic template suggests that acts that directly involve suffering are more likely to be seen as "typical" moral violations. Indeed, as many others have documented, concerns about harm are universal, emerging across countries, cultures, and political orientations, whereas concerns about authority and purity appear more limited (Graham et al., 2009; Haidt, Koller, & Dias, 1993). The presence or absence of harm also distinguishes moral transgressions from conventional transgressions (Nucci, Turiel, & Encarnacion-Gawrych, 1983; Turiel, 1983). Furthermore, concerns about harm emerge remarkably early in development (Blair, 1995; Hamlin, Wynn, & Bloom, 2007) and can be seen even in nonhuman animals (de Waal, 2006)—even rats respond to the suffering of conspecifics (Church, 1959).

One simple experimental method for determining the typicality of examples within a concept is to examine accessibility (Mervis & Rosch, 1981), and one measure of accessibility is spontaneous recall (Bargh & Thein, 1985). If interpersonal harm is the essence of morality, then asking people to think of an act that is morally wrong should prompt recall of initial examples of direct suffering. This study was conducted by asking approximately 100 participants from diverse cultures to "list an act that is morally wrong" (Gray & Ward,

2011). The majority of participants (51%) listed murder/killing/raping/intentionally harming another—all acts of direct harm. Other acts listed included stealing (19%), and adultery (7%)—both of which cause harm—and cheating/lying (10%). The combination of homosexuality, bribery, nepotism, gossip, having sex in public, and betraying your siblings all accounted for less than 10%.

A dyadic model of morality predicts that, when multiple moral concerns are in conflict, harm should trump other concerns. Van Leeuwen and Park (2011) tested this directly by asking participants to select the moral concern most important for building an ideal society. Although participants could select among all of Haidt's five moral domains (Haidt & Graham, 2007), the most commonly selected domain was harm (approximately 50% of participants), regardless of the participant's political orientation. That conservatives preferred to address harm-related concerns is even more striking given that the task was to build an ideal society, where group-related concerns might dominate. More evidence for the dominance of harm-related concerns comes from Wright and Baril (in press), who demonstrate that conservatives fundamentally possess a harm-based morality: Under cognitive load, conservatives deemphasize the domains of authority, ingroup, and purity, suggesting that mental effort is required to moralize domains that lack a clear dyadic structure.

Extensive evidence suggests not only that moral acts can be defined in terms of intention and suffering but also that perceptions of suffering unify various moral domains, and that harm is the most important of moral domains. The importance of suffering in morality—and the explanatory power of the moral dyad—seems to be a general rule, but next we review apparent exceptions.

Reconciling Apparent Exceptions With Dyadic Morality

We suggest that morality can be understood through the lens of interpersonal harm—the combination of intention and suffering. Yet there are instances that appear to defy this dyadic structure. We explore three potential counterexamples and reconcile them with dyadic morality.

Honor Killing

The moral decisions of other cultures can come as a shock to liberal Westerners, as they appear to ignore the presence of suffering victims. Rai and Fiske (2011) described the phenomenon of honor killing, in which a rape victim is subsequently murdered by her family, who then celebrate her death. But consider the following facts: In rural India and Pakistan, marriage is more of an economic agreement between two families, in which wives are exchanged for dowries—but only if the woman is seen as pure (Husseini, 2009). A bride losing her virginity before marriage not only damages her spiritually but also threatens the groom with possible sexually transmitted diseases. Because she is now unmarriageable, this woman will continue to consume her family's resources, damaging the family's economic well-being. Similar cost–benefit explanations can be seen in infanticide of ancient Sparta (Patterson, 1985), and in contemporary India (Sen, 2002) and China (Croll, 2000), where the relatively lower status and earning potential of women prompts some parents to kill female babies.

The killing of individuals in favor of familial benefit is facilitated by stripping away mind from victims (i.e., dehumanization, as reviewed earlier; Haslam, 2006; Waytz & Young, in press). Indeed, an exposé about Yemeni child brides (Gorney, 2011), where girls as young as 5 are married to older men, revealed that those who facilitate the marriages fail to perceive any suffering in the young victims. Thus, these culturally motivated killings are noteworthy not because they devalue suffering per se, but because of the extent to which people are stripped of mind to justify potential collective benefits, a phenomenon also found in the West (e.g., forced sterilization [Zigler, 1967] and the Tuskegee medical trials [Reverby, 2000]). Finally, honor killings and child marriages are not uniformly or even typically celebrated; many within those cultures and families revile the obvious suffering they cause (most notably other women; Gorney, 2011; Husseini, 2009).

Moral Dumbfounding

One argument against our account of moral judgment is moral dumbfounding, in which people are left speechless after harm-based explanations for wrongness have been nullified (Haidt, 2001). For example, people continue to reject sibling incest even when both parties use protection and enjoy doing it. Just because harm has been explicitly nullified, however, does not preclude implicit perceptions of harm from driving moral judgment. For example, a person standing on the Grand Canyon Skywalk, a walkway of transparent glass cantilevered over the Grand Canyon, may still be terrified even though she knows she is safe. This does not imply that she must be afraid of something else, but rather that her fear of heights defies explicit knowledge (Gendler, 2008a, 2008b). In the case of moral dumbfounding, people may still be reacting to the perceived harmfulness of transgressions despite explicit denials of harm. Indeed, such a reliance on intuitive perceptions of harm is consistent with the social intuitionist model (Haidt, 2001); we simply suggest that the intuition of harm is naturally tied to judgments of wrongness.

Disgust-Related Judgments

One set of moral judgments that appears to rely less upon perceptions of mind than mere intuition are those that involve purity and disgust. Judgments of wrongness are typically linked to feelings of disgust (Chapman, Kim, Susskind, & Anderson, 2009; Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005), an aversive emotion focused on threats of contamination to the body (i.e., tainted meat, poor hygiene, and body envelope violations; Rozin, Haidt, & McCauley, 2008). The bodily focus that disgust induces may serve to reduce the role of mind perception in moral judgments and inhibit the perception of mind more generally. After all, people are intuitive dualists, perceiving minds and bodies as distinct (Bering, 2006; Bloom, 2004; Demertzi et al., 2009) and so stimuli or emotions that induce a bodily focus reduce ascriptions of mental states (Archer, Iritani, Kimes, & Barrios, 1983; Heflick & Goldenberg, 2009; Nussbaum, 1995). For example, making the bodies of women salient can increase dementalized perceptions of them (Cikara et al., 2010; Heflick & Goldenberg, 2009; Loughnan, Haslam, Murnane, et al., 2010), as can focusing on the bodies of medical patients (K. Gray, Knickman, et al., in press). Research on stereotyping also reveals a link between the experience of disgust and reduced mind perception (Harris & Fiske, 2006). Furthermore, Young and Saxe (in press) showed that the perceived wrongness of bodily disgusting acts (e.g.. incest) depends less on mental state dimensions (e.g., intent).

Although the link between disgust and wrongness appears to be unrelated to intention, disgust initially evolved to protect people from bodily harm (Rozin et al., 2008), and so the experience of moral disgust can be seen as a heuristic for potential suffering. Although there are cases where eating roadkill is safe, a general aversion to carrion is adaptive, and a moral aversion
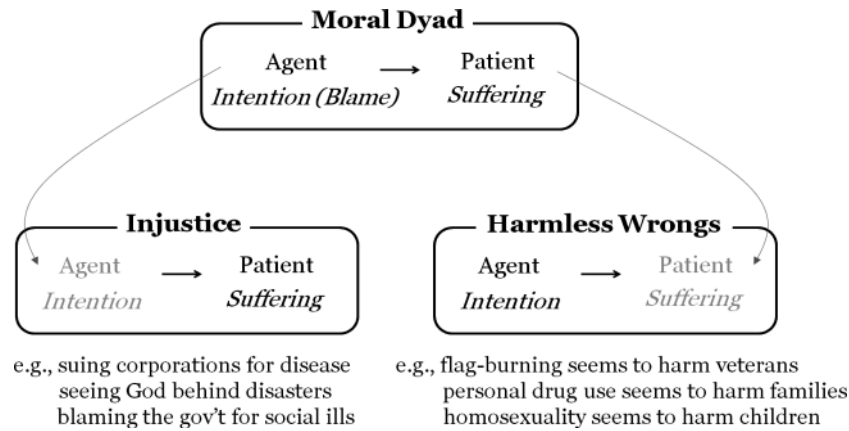
Figure 6. Dyadic completion. *Note.* A dyadic template compels people to see a blameworthy agent for unjust suffering and to see immoral acts as inducing harm.

to such acts can powerfully motivate behavioral avoidance (Gigerenzer & Todd, 1999). In addition, priming harm renders disgusting bodily related transgression more immoral (Rosenblatt et al., 1989), whereas other studies find that simply focusing on the body can sharpen the salience of suffering (Gray, Knobe, Sheskin, Bloom, & Barrett, 2011). Finally, research finds that disgusting but ostensibly harmless moral violations are linked to perceived harm (Royzman, Leeman, & Baron, 2009), suggesting that the dyad may function even in the context of disgust.

Reconciling Mind Perception and Different Moral Domains

A dyadic template suggests that people understand morality as a combination of agent and patient, intention and suffering. This does not mean, however, that there are not *descriptively* different domains of morality; that conservatives see some issues as morally relevant, whereas liberals see them as personal choice is clear. Our point is that these domains are not irreconcilable with each other—they can be linked through mind perception. Next we explore how dyadic morality can also account for two novel phenomena in moral psychology—dyadic completion and moral typecasting.

Novel Phenomena of Dyadic Morality

The Dyadic Structure of Morality Both Compels and Constrains Judgments of Mind and Morality

Mind perception and dyadic morality dovetail with a variety of other moral theories, and also have the power to highlight novel phenomena concerning morality and mind. We explore two such phenomena that the moral dyad explains—dyadic completion and moral typecast-

ing. The first compels judgments of mind, and the second constrains judgments of mind and morality.

Dyadic Completion

If our template of morality is dyadic, we should be compelled to complete the moral dyad when it appears incomplete. This dyadic completion can occur in two complementary ways. First, when we see someone blameworthy—an apparent moral agent—we should complete the dyad by inferring the presence of another mind to suffer—a moral patient. Second, when we see a suffering patient, we infer the presence of another mind to take responsibility as a moral agent (Figure 6). We suggest the phenomenon of dyadic completion occurs at an intuitive level—like the Gestalt completion of the Kanizsa triangle (Figure 4).

The link from agent to patient—seeing suffering in response to blameworthy intention—has been covered in previous sections. Recent research suggests that perceptions of intention and blame are also translated into increased suffering in physical experience. In one study, participants received electric shocks that were administered either intentionally or accidentally, and though the shocks were identical in voltage, the more intentional (and blameworthy) shocks were experienced as physically more painful (Gray & Wegner, 2008). This increased experience of pain from intentional shocks also translates into increased skin conductance responses (Gray, 2010a).[4]

A dyadic moral template should lead to dyadic completion in the reverse direction; otherwise inexplicable suffering should prompt perceptions of moral agency. Although good events can prompt such attributions

---

[4]Extending these findings to the domain of good deeds, one study revealed that praiseworthy intention increases pleasure, as intentionally administered massages were rated as more pleasurable than identical unintentionally administered massages (Gray, in press).

(e.g., Pepitone & Saffiotti, 1997; Spilka & Schmidt, 1983), bad events are psychologically more powerful than good events, and so we would expect suffering to lead to increased perceptions of moral agency (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Pennebaker & Beall, 1986; Taylor, 1991). Indeed, one study finds that those who receive unfair splits on a dictator game are more likely to perceive the hand of an intentional agent (Morewedge, 2009). The "Knobe effect" (Knobe, 2003) is a similar phenomenon, whereby people rate blameworthy side effects as more intentional than praiseworthy side effects.

Of interest, when suffering cannot be attributed to human agents, people often blame nonhuman agents. For example, in medieval France, failed harvests and terrible accidents were sometimes ascribed to animals, which were tried in the local legal system (Humphrey, 2003). In one case, a pig was discovered next to a dead child and was subsequently tried, found guilty, and hanged (Oldridge, 2004). More typically, tragedy is ascribed to supernatural agents, such as God, gods, or malevolent spirits (Boyer, 2001; Bulman & Wortman, 1977; Gall, 2004; Gray & Wegner, 2010a; Kay, Gaucher, McGregor, & Nash, 2010; Lewis, 1995; Pargament et al., 1990; Spilka, Zwartjes, & Zwartjes, 1991). Anthropologists have documented many cases in tribal societies where deaths and illnesses are ascribed to spirits and witchcraft (Boyer, 2001; Lewis, 1995), and a recent study has even linked suffering to the belief in God in the United States (Gray & Wegner, 2010a). In this study, the amount of suffering in each state (as indicated by lower scores on the United Health Foundation's health index) correlated with the per-

centage of people in each state who strongly believe in God. Such dyadic completion also occurs in response to subtle cues, where simply framing someone as a victim makes nearby others appear more like moral agents (Gray & Wegner, 2009; Young & Phillips, 2011).

People appear to be compelled to complete the moral dyad, seeing suffering in response to blameworthy intention, and seeing blameworthy intent in response to suffering. Next, we describe how the moral dyad can constrain perceptions of others.

Moral Typecasting

Just as moral acts may be defined by mind perception, the minds of others are also defined by their moral acts. A dyadic template of morality suggests that people are categorized as either moral agents or moral patients—a phenomenon called moral typecasting.

The word "typecasting" has its roots in Hollywood, and one enduring example of such typecasting is Leonard Nimoy, best known for his role as *Star Trek*'s Spock. Although he brought much to the role, the role itself influenced how Nimoy was perceived. First, people assumed Nimoy must be as rational as his character; in real life, Nimoy could have been passionate and erratic, but his Vulcan role led people to see him otherwise. Second, his role as Spock forever defined him; despite the variety of other characters he attempted, people could not help but see him as anything other than Spock. In fact, Nimoy titled his initial 1977 autobiography, *I Am Not Spock*, but by 1995 he resigned himself to this typecasting, titling
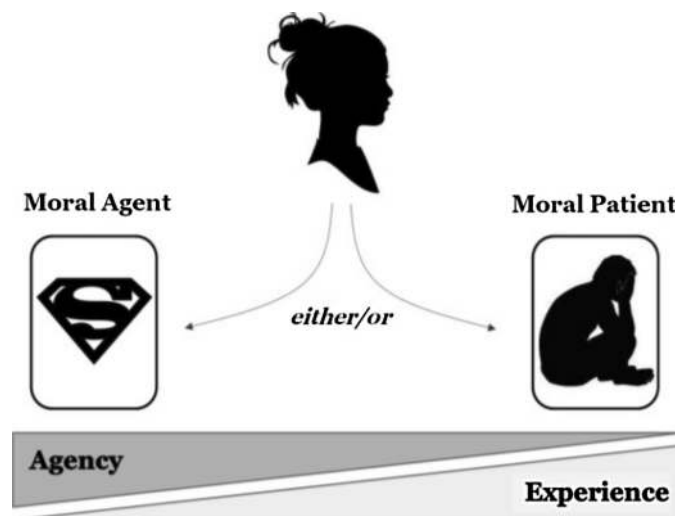

Figure 7. Moral typecasting. *Note.* Those cast as moral agents are seen to have agency, whereas those cast as moral patients are seen to have experience. In addition, people are generally seen to be either moral agents or moral patients, making the normally orthogonal dimensions of agency and experience inversely related.

his second biography, *I Am Spock*. Just as we typecast actors, we also typecast the people around us into enduring moral roles, preventing them from taking on other moral roles.

Moral typecasting also influences our perception of the target person's mind. When someone is categorized as a moral agent, observers automatically infer the capacity for agency. This means that simply doing something good or evil can bring along corresponding attributions of intention (especially evil, Knobe, 2003; see Gray & Wegner, 2009). Likewise, when someone is categorized as a moral patient, people automatically infer the capacity for experience (Figure 7) and greater sensitivity to pain (Gray & Wegner, 2009). The link between moral role and mind can also extend beyond mere perception; one study found that thinking of yourself as a hero or a villain actually increases physical agency, as measured by the length of time a weight could be held (Gray, 2010b).

Typecasting further suggests that people are cast into enduring and mutually exclusive moral roles—as *either* moral agents *or* moral patients. Those who are moral agents are seen to be incapable of being a moral patient; those who are moral patients are seen to be incapable of being an agent (Figure 7). Although the two-dimensional structure of mind perception suggests that perceptions of agency and experience are independent, within a moral context, perceptions of moral agency and moral patiency may oppose each other.

Think of a typical moral misdeed such as theft, in which one person (the agent) steals money from another (the patient). Now imagine that both the thief and the victim are the same person—the act loses its moral status and becomes simply taking money out of your own wallet. Moral acts therefore typically require two different people; the agent cannot be the patient, and the patient cannot be the agent. One apparent exception to this rule is suicide, but still people perceive victims (the remaining family) or perpetrators (those who drove the person to suicide). Another apparent exception is consensual incest, but observers often apply a dyadic template and place one person as agent and one as patient, such as in a publicized case of a father (agent) sleeping with his adult daughter (patient; Tsoulis-Reay, 2010).

In general, then, a dyadic template splits moral acts into two different and asymmetric roles, a structure also found in other social domains (Baldwin, 1992; Wegner & Vallacher, 1977). For instance, in dominance relations, (a) there needs to be at least two people so that one can exert power over another, and (b) if person A exerts power over person B, it implies that person B cannot exert power over person A (De Soto, 1960). Moral typecasting is the idea that this *either/or* of moral acts criterion applies more broadly to people: People are generally seen as either moral agents or moral patients.

If people are seen as either heroes and villains or victims and beneficiaries, then it should be difficult to see heroes and villains as capable of suffering. Indeed, both good and bad moral agents are perceived to feel less pain from injuries (Gray & Wegner, 2009). Typecasting also suggests that it should be difficult to see suffering victims as blameworthy villains. Accordingly, framing yourself as a victim is a more effective way of escaping blame than framing yourself as a hero (Gray & Wegner, 2011b; Weiner, 1980); although heroes may have good deeds to their credit, they still remain moral agents and therefore can be assigned more praise *and* more blame. Simply perceiving someone in terms of their experience (e.g., fear, hunger, rage) can also reduce blame (Jenkins & Wegner, 2011). For example, people excuse crimes of passion, in which the agents are victims of their own emotions (Finkel & Parrott, 2006); it is difficult to reconcile rage and fear with the intention and planning that typifies moral agency.

The blame-reducing effects of moral patiency also appear to apply to perceptions of the self, such that people who are made to feel like victims, act more immorally, perhaps because they feel incapable of earning blame (Zitek, Jordan, Monin, & Leach, 2010). Turning yourself into a victim after committing a transgression also reduces personal guilt: Allowing people to shock themselves after antisocial actions made people feel better (Bastian, Jetten, & Fasoli, 2011). The same sentiment may also apply when we punish others, helping to make sense of people's lust for retribution (Carlsmith, 2006; Carlsmith & Sood, 2009): Pain transforms offenders from agents to patients, redeeming them in the eyes of society.

Typecasting can have some surprising effects. For instance, the apparent insensitivity of moral agents to pain leads people to endorse harming not only villains but also heroes, whose past good deeds should earn them reward instead of punishment (Gray & Wegner, 2009; see also Monin, Sawyer, & Marquez, 2008). This finding contradicts the belief in a just world (Lerner, 1980), as does the finding that victims receive less blame than both do-gooders and the average person (Gray & Wegner, 2011b). Numerous studies make clear that people are only too willing to blame the victim (Furnham, 2003; Janoff-Bulman, Timko, & Carli, 1985; Lerner & Miller, 1978; Lerner & Simmons, 1966), but these studies frequently involve some kind of uncomfortable arousal or complicity on the part of participants (Cialdini, Kenrick, & Hoerig, 1976). For example, one study found that participants blamed the victim when they were responsible for the suffering of the victim, but not otherwise (Cialdini et al., 1976). A more recent study found that those uninvolved in the torture blamed victims *less* when they suffered more, as typecasting predicts (Gray & Wegner, 2010b).

Moral typecasting suggests that, within morality, perceptions of the dual dimensions of mind are not

independent. Instead, people view others as moral Necker cubes, as either agents or patients, capable of either intention and blame, or experience and pain. This either/or perception stems from the structure of the moral dyad. Other research even suggests that this phenomenon extends more broadly, to perceptions of entire groups of people like countries (Kervyn, Yzerbyt, Demoulin, & Judd, 2008).

In sum, dyadic morality can uniquely explain both dyadic completion and moral typecasting. Next we explore how mind perception may help to unify morality across levels of analysis.

Mind Perception and Other Formulations of Morality

Mind Perception Can Unify Morality Across Different Levels

Many phenomena can be understood on different levels. The concept of "university" could be understood as a set of buildings, a collection of individual students and professors, or a broader set of cultural values. Some definitions, however, transcend these levels of description: Universities are broadly about learning and research. Buildings are where research takes place, professors direct the research, and an emphasis on learning shapes a university's cultural values. Morality can also be understood at multiple levels, but we suggest that mind perception provides a unified understanding of moral judgment. In this section, we divide a number of moral theories into three different levels—group, individual, and intrapersonal—and explore how each level can be reframed in terms of mind.

First, at the group level, morality concerns *community*—how people navigate group living and standards (Graham et al., 2009; Rai & Fiske, 2011; Shweder et al., 1987). Second, at the level of the individual, morality concerns *character*—how we judge not specific acts but the agents who perform them (Alicke, 2000; Pizarro & Tannenbaum, 2011; Tannenbaum, Uhlmann, & Diermeier, in press). Third, within the individual, moral judgments consists of combining affect and cognition, which in turn depend on component principles or grammatical breakdowns of moral acts. We refer to this view as the *componential* view (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Hauser, 2006; Mikhail, 2007; Nichols, 2002).

Community

Two major theories of morality emphasize the community level. First, moral foundations theory (Graham & Haidt, 2010; Haidt, 2007) emphasizes cultural differences in morality. This theory suggests that different cultures select different moral principles depending upon ideology and religion. On this view, cultures and communities build narratives around different moral domains. A second theory, relationship regulation theory (Rai & Fiske, 2011), suggests that distinct motives for maintaining different social relationships determine whether an action is considered right or wrong. This theory suggests that the moral character of any action depends on the specific relationship between people.

Mind perception is essential to a community-based view of morality. Mind perception forms the basis for cooperation, coordination, and communication necessary for building and maintaining social groups (Baron-Cohen, 1995; Epley & Waytz, 2009; Humphrey, 1976; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Indeed, theories suggest that the evolution of mind perception was driven by the same concerns for the evolution of morality, allowing individuals to navigate group living (DeScioli & Kurzban, 2009; Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007).

Not only do perceptions of individual minds help facilitate binding individuals into groups, but groups themselves are also perceived to have mind (Knobe & Prinz, 2008; Knobe, Jenkins, Dodell-Feder, & Saxe, 2011; Waytz & Young, in press). Perceptions of group mind can help explain moral behavior such as self-sacrifice for one's country and religion (Routledge & Arndt, 2008), and also other cases of putting group concerns before individual interests (Ridley, 1998).

Group formation and the survival of specific cultures in a competitive evolutionary landscape are also supported by perceiving mind in supernatural agents. There is significantly more temptation to act selfishly when alone, but one is never alone if one perceives a supernatural agent who monitors all actions. Studies indicate that being primed with God leads people to cheat less (Bering, McLeod, & Shackelford, 2005) and to be more generous in social and economic exchanges (Shariff & Norenzayan, 2007). The problem of selfish behavior is even more pressing in large societies, where anonymity overcomes concerns about reputation (Nowak, 2006), but again mind perception comes to the rescue of the group. Cooperation in large societies is encouraged by people perceiving God to have a specific kind of mind (i.e., punishing; Norenzayan & Shariff, 2008). Thus, mind perception not only underlies judgments of individual moral acts but also helps solve the problem of how groups evolve in the first place.

Character

The character view of morality suggests that people base their moral judgments not on the quality of a particular action but on whether they deem the *actor* to be a good or bad person (Pizarro & Tannenbaum, 2011; Tannenbaum et al., in press). In support of this view, the same action completed by someone with a bad (vs. good) character is rated more harshly (Alicke,

1992), and ostensibly less harmful acts (cat beating; Tannenbaum et al., in press) can be rated as worse than more harmful acts (wife beating) when they suggest an evil character.

Mind perception provides a basis for a character-centered view of morality, because mind perception forms the basis of person perception. Theories of behavioral attribution dating back to Heider (1958) suggest that inferences of a person's traits from their behavior center on the inference of a person's intentions. Instead of describing causal attributions of behavior in terms of the dispositional-situational dichotomy that would come to dominate the field of social psychology, Heider (1958; see also Jones & Davis, 1966) described this dichotomy in terms of personal causality (actions caused by intention) versus impersonal causality (actions caused unintentionally). Modern theories of dispositional inference also suggest that the inference of people's beliefs and motives underlies trait attributions to these people (Malle, 2006, 2011; Reeder, 2009). Indeed, the attribution of mental states appears to occur more quickly and automatically than the attribution of traits (Van Overwalle, Van Duynslaeger, Coomans, & Timmermans, in press). Thus, character assessments stem from assessments of mind and intention—not necessarily for the act in question but across a variety of previous acts.

Componential

The componential view of morality suggests that moral judgments are driven by two core components—affective reactions and cognitive processes (Greene, 2007; Nichols, 2002) that can be further dissected into intuitive principles or parameters. One version of this view suggests that moral judgment reflects a "universal moral grammar," which parallels the deep structure of language (Hauser, 2006; Mikhail, 2007; Rawls, 1971). In particular, proponents of this "linguistic analogy" argue that universal moral grammar consists of sets of rules that take a wide array of nonmoral inputs (e.g., actions, causes, emotion, perceived intentions) and translates them into moral judgments (Cushman & Young, in press; Mikhail, 2007). These computations can be simple, "ME HURT YOU = WRONG" (Greene et al., 2004), or more complex, "INTENT + CAUSE + HARM = WRONG" (Mikhail, 2007).

We suggest that mind perception is crucial for switching on the "moral faculty." Factors of intent, cause, personal force, and valuation may be combined into a moral judgment, but mind perception precedes these computations. For example, assessments of cause are relevant only insofar as the cause is an agent, with the relevant kind of mind (Muentener & Lakusta, 2011). Assessments of intent require first establishing that a mind exists before specific mental contents can be inferred. Assessments of personal force—the means by which intentional harm was caused (Greene, 2007)—also depend on the presence of an agent with a mind to power the act. Finally, how much harm only matters, once again, if the harm is caused by a mindful agent, not, for instance, a thunderstorm or a pack of wolves (Nichols, 2002). In other words, our moral code may forbid acting intentionally, with one's own personal force, to cause a great deal of harm. But how—and how much—harm is done matters only when done by an agent toward a patient.

Recent theorizing highlights the role of affect as a component of moral judgment (Greene, 2007; Nichols, 2002), but what elicits this affect? Most often, it seems to be triggered by perceiving a mind—by the outrage of intentional harm (Kahneman et al., 1998) or by the aversion to suffering (Blair, 1995; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). Moral judgments may be underlain by components of "cognition" and affect, but both appear linked to mind perception.

Unification

Distinct theories of morality have in common one key component: mind perception. On the community view, immoral actions undermine group cohesion or specific relationships within the group. Assessing social relationships requires assessing the minds of those individuals in relation—who are they, how are they related, and whether they know what they're doing. On the character view, people, not actions, are branded immoral. Evaluating a person's character depends crucially on knowing what's on his or her mind—external actions don't provide enough insight into the "deep self" (Frankfurt, 1971; Sripada, 2009). On the componential view, moral judgments are driven by principles or parameters that hinge upon the perception of other minds, whether they involve perceptions of intent or harm.

Development of the Moral Dyad

A dyadic template appears to explain adult moral judgments, but how is such a template acquired? We suggest this template builds on three ontogenetically early components: (a) an understanding of causation, (b) perceptions of intention and pain, and (c) the empathic aversion of others' pain.

Whether we observe billiard balls or cars colliding, the human mind understands instances of physical causation in terms of an agent (cause) and a patient (effect). This dyadic understanding of physical causality emerges at a very young age (Rochat, Striano, & Morgan, 2004), as does the ability to perceive suffering in others (Blair, 1995; Sagi & Hoffman, 1976)

and the ascription of intention (Gergely, Nádasdy, Csibra, & Bíró, 1995; Saxe, Tenenbaum, & Carey, 2005). Combining a dyadic conception of physical causation and ascriptions of intention and suffering provides a template of *mental* causation in which one person's intention causes another's person's pain.

Indeed, there are many reasons why intention may be perceived as causally tied to suffering. First, causes and effects seem to be most easily understood when occurring at the same explanatory level (i.e., the level of mind; Davidson, 1970). Second, the same mental events (e.g., pain) can be arrived at via different physical means (e.g., kicking, hitting, social exclusion), and mental causes can account for all of these means as one broad causal structure (i.e., she means to harm me; Lombrozo, 2010). Third, pain and pleasure are very important mental events to understand, and important events are more likely to be attributed to intentional agents (Epley, Waytz, & Cacioppo, 2007); this "intentional stance" allows for a more powerful route of prediction and control (Dennett, 1989).

It seems likely that people tie together intention and suffering into a dyadic structure of agent and patient (Brown & Fish, 1983), but not all instances of causation between minds is moral. How does this general causal template develop into the *moral* template? More specifically, how does pain get linked not just with intention but also with blame? Quite simply, ascriptions of blame can stem from the aversiveness of experienced pain or the aversiveness of pain perceived in the minds others, via empathic aversion (Blair, 1995; Davis, 1996; Decety, 2011; Preston & de Waal, 2001; Singer et al., 2004). Even newborns find the distress of others aversive, providing a ready route through which the intentional causation of pain can be seen as blameworthy (Martin & Clark, 1982; Sagi & Hoffman, 1976). Indeed, Blair (1995) suggested that this innate aversion to others' pain underlies the general acquisition of morality. Research with infants also supports this idea—even 6-month-olds pass judgment on those who harm others (Hamlin et al., 2007). We therefore suggest that empathic aversion is the key for turning the dyad into the moral dyad at an early developmental stage (Nichols, 2002).

Of course, empathic aversion alone does not make an act immoral, because both hiking accidents and assault may invoke empathic responses, but only the later is judged as immoral. Nichols (2002) suggested that we need norms to distinguish between the immoral and the simply unfortunate; his idea is that only acts that both generate negative affect and violate norms are judged as immoral. Norms are important in structuring the moral world, as is negative affect, but these factors are not sufficient to account for moral judgment. For example, if a child wears pajamas to school when no one else does, she not only violates norms but also would likely feel terrible. Nevertheless,

this act is not viewed as immoral (Turiel, 1983). We suggest that immoral acts are norm violations that match a dyadic template: Acts are wrong when they involve the intentional causation of suffering. In other words, empathic aversion is translated to immorality when pain is caused by an intentional agent (see also Royzman, Leeman, & Baron, 2009).

The acquisition of a dyadic template is not mysterious. Babies are born ready to apply a causal structure of agent and patient to events, and attribute agency to those who cause events. Pain is one mental event that requires mental explanation, and its aversiveness helps turn mental causation into moral causation. Once built, this dyadic template of intention and suffering serves as a way to understand moral acts in general.

Predictions

Theories should not only account for previous findings but also generate novel predictions. Here, we present two areas of future research—individual differences in morality and mind perception and the difference between good and bad moral acts.

Individual Differences

There is no doubt that different people think that different acts constitute immoral behavior. Conservatives believe that having sex with dead chickens is morally wrong, whereas liberals believe it to be a matter of personal choice (Graham et al., 2009). These individual differences lead to a number of important phenomena—especially political disagreements (Ditto & Koleva, 2011). As we suggest, moral judgments are rooted in mind perception; thus, future research should reveal corresponding differences in mind perception. Studies suggest that people do vary in their ascriptions of mind (Gray, Jenkins, et al., 2011; Waytz, Cacioppo, & Epley, 2010), and these individual differences may be linked to political orientation. Conservatives appear to see both more agency and experience in other people (Gray, Knobe, Sheskin, Bloom, & Barrett, 2011), which can explain how conservatives see both more nefarious intention and complementary suffering in the world. Furthermore, relative to liberals, conservatives ascribe more mind and moral rights to entities such as fetuses and vegetative patients (Gray et al., 2007).

In other cases, though, liberals may ascribe relatively more mind. Liberals are generally more concerned about the environment and animal rights, and vegetarians (usually liberals) are correspondingly more likely to ascribe mind to farm animals (Bastian, Loughnan, Haslam, & Radke, 2011; Loughnan, Haslam, & Bastian, 2010). It should also be the case that antiwhaling activists will ascribe more mind to marine mammals, and antiwhaling activists will ascribe more mind to "Mother Earth."

Dyadic morality may also motivate a new model of moral character—the characteristic ways in which people react to moral situations. Dyadic morality suggests that people divide the moral world into four moral characters—heroes, villains, victims, and beneficiaries—and such self-perceptions may translate into the behavior of people themselves. In other words, one person may characteristically be a victim and typically feel harmed by others (Janoff-Bulman, 1979), whereas someone else may be characteristically a hero and typically feel that he/she is helping others (Walker, Frimer, & Dunlop, 2010). These self-perceptions may also translate into physical effects, whereby those who see themselves as agents may actually be able to exert more physical self-control (Gray, 2010b).

Good Versus Bad

Moral judgments may differ not only between people, but the same person may also judge good and evil acts differently. It may be intuitively appealing to think of moral deeds as simply the opposite of immoral deeds, but research paints a more complex picture. For example, good deeds have reduced affective power than bad deeds (Baumeister et al., 2001) and are typically less likely to motivate sense-making processes (Taylor, 1991; Ybarra, 2002). Perhaps most relevant to the link between mind perception and morality, good deeds are less likely than bad deeds to prompt perceptions of intentionality, likely because they are consistent with situational constraints (Phillips & Knobe, 2009; Ybarra, 2002) and our lay-theories of character (Pizarro, Uhlmann, & Salovey, 2003). Nevertheless, we suggest that a dyadic template underlies our understanding of good deeds. For example, studies on moral typecasting uncover similar attributions of agency to both good and evil moral agents, and similar propensities to ascribe pain to both heroes and villains (Gray & Wegner, 2009). Future studies should test whether a dyadic template underlies different conceptions of virtue. Just as perceived harm can unite various dif-

ferent moral domains, so should perceived help unite different domains of goodness.

One promising approach for comparing the effects of good versus bad is provided by Janoff-Bulman and colleagues (Janoff-Bulman, 2009; Janoff-Bulman, Sheikh, & Hepp, 2009), who outlined individual differences in the psychological weight of good versus bad deeds. Specifically, they found that liberals focus on engaging in good behaviors (prescriptive morality), whereas conservatives focus on refraining from bad behaviors (proscriptive morality). These differences are tied to fundamental motivational orientations of approach (liberals) and avoidance (conservatives).

Some recent research has tried to integrate the good/bad distinction with agent/patient distinction in terms of a two-dimensional space of moral emotion (Gray & Wegner, 2011a; see also Haidt, 2003). This space yields four quadrants of emotions that are felt toward each of four different moral characters, as previously outlined: heroes, villains, victims, and beneficiaries (Figure 8). For example, the emotions felt toward heroes (in the help/agent) are inspiration and elevation (Algoe & Haidt, 2009), whereas those felt toward victims (in the harm/patient) are sympathy and sadness.

In addition to providing a conceptual space for understanding moral emotions, this model makes specific testable predictions. The first prediction is that emotions in the same quadrant should reinforce each other. For instance, feeling anger toward someone in a moral situation should predispose you to feel disgust; indeed, people typically feel both disgust and anger toward villains such as violent racists (Haidt, Rozin, Mccauley, & Imada, 1997). The second hypothesis—suggested by moral typecasting—is that emotions in different quadrants should conflict with each other. If people are seen as either moral agents or moral patients, then the more sympathy is felt toward someone, for example, the less anger should be felt toward them. Studies by Weiner and colleagues suggest this to be the case (Schmidt & Weiner, 1988; Weiner, 1980). The

Figure 8. Linking moral emotions to the dimensions provided by dyadic morality. *Note.* Reprinted with permission from Sage. *Source.* Gray and Wegner (2011a).

third hypothesis—suggested by dyadic completion—is that agent-related emotions should potentiate patient-related emotions toward *another* person, and vice versa. For example, if you feel anger and disgust toward someone, you should be potentiated to feel sympathy and sadness toward another person. Of course, the moral emotional world is more complex that just these four quadrants, and includes other emotions like gratitude (Bartlett & DeSteno, 2006), jealousy and pride (Valdesolo & DeSteno, 2011), guilt and shame (Keltner & Buswell, 1996), and mirth (Strohminger, Lewis, & Meyer, 2011), but this model provides testable hypotheses linked to dyadic morality.

Conclusion

We have suggested that mind perception is the essence of morality and that moral judgments are rooted in a dyadic template of two perceived minds—an agent and a patient. Dyadic morality is suggested by the correspondence between mind perception and morality and the enduring presence of perceived suffering in moral transgressions. It not only accounts for diverse findings in moral psychology but also explains the phenomenon of moral typecasting and dyadic completion. Decades ago, Picasso captured the essence of bulls. Although his elegance could never be matched, we have attempted to follow his lead and capture the essence of morality—not with lines—but with minds.

Acknowledgments

Note

Address correspondence to Kurt Gray, Department of Psychology, University of North Carolina, CB #3270 Davie Hall, Chapel Hill, NC 27599. E-mail: kurtjgray@gmail.com

References

Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: The "other-praising" emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology, 4*, 105–127. doi:10.1080/17439760802650519

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology, 63*, 368–378.

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*, 556–574.

American Law Institute. (1962). *Model Penal Code.*

Anderson, S. W., Barrash, J., Bechara, A., & Tranel, D. (2006). Impairments of emotion and real-world complex behavior following childhood- or adult-onset damage to ventromedial prefrontal cortex. *Journal of the International Neuropsychological Society, 12*, 224–235. doi:10.1017/S1355617706060346

Archer, D., Iritani, B., Kimes, D. D., & Barrios, M. (1983). Face-ism: Five studies of sex differences in facial prominence. *Journal of Personality and Social Psychology, 45*, 725–735. doi:10.1037/0022-3514.45.4.725

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

Aristotle. (2009). *Nicomachean ethics* (W. D. Ross, Trans.). New York, NY: World Library Classics.

Baird, J. A., & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development, 2004*, 37–49. doi:10.1002/cd.96

Baird, J., & Moses, L. (2001). Do preschoolers appreciate that identical actions may be motivated by different intentions? *Journal of Cognition and Development, 2*, 413–448. doi:10.1207/S15327647JCD0204_4

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin, 112*, 461–484. doi:10.1037/0033-2909.112.3.461

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*, 364–374.

Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology, 49*, 1129–1146. doi:10.1037/0022-3514.49.5.1129

Baron-Cohen, S. (1995). *Mindblindness.* Cambridge, MA: MIT Press.

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

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17. doi:10.1023/A:1005653411471

Barrash, J., Tranel, D., & Anderson, S. W. (2000). Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. *Developmental Neuropsychology, 18*, 355–381.

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasurement of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition, 121*, 154–161.

Bartlett, M. Y., & DeSteno, D. (2006). Gratitude and prosocial behavior: helping when it costs you. *Psychological Science: A Journal of the American Psychological Society / APS, 17*, 319–325. doi:10.1111/j.1467-9280.2006.01705.x

Bastian, B., Jetten, J., & Fasoli, F. (2011). Cleansing the soul by hurting the flesh. *Psychological Science, 22*, 334–335. doi:

Bastian, B., Laham, S. M., Wilson, S., Haslam, N., & Koval, P. (in press). Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology*. doi:10.1348/014466610×521383

Bastian, B., Loughnan, S., Haslam, N., & Radke, H. R. M. (2011). Don't mind meat? The denial of mind to animals used for human consumption. *Personality and Social Psychology Bulletin, 38*(2), 247–256. doi:10.1177/0146167211424291

Baumeister, R. F. (1999). *Evil: Inside human violence and cruelty.* New York, NY: Holt.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370. doi:10.1037/1089-2680.5.4.323

Benson, M., & Kugler, J. (1998). Power parity, democracy, and the severity of internal violence. *The Journal of Conflict Resolution, 42*, 196–209.

Bering, J. M. (2006). The folk psychology of souls. *Behavioral and Brain Sciences, 29*, 453–462.

Bering, J., McLeod, K., & Shackelford, T. (2005). Reasoning about dead agents reveals possible adaptive trends. *Human Nature, 16*, 360–381. doi:10.1007/s12110-005-1015-2.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition, 57*(1), 1–29. doi:10.1016/0010-0277(95)00676-P

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*, 698–718. doi:10.1016/j.concog.2005.06.004

Bliss-Moreau, E., & Barrett, L. F. (2009). Affect as a psychological primitive. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 167–218). New York, NY: Academic Press.

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

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York, NY: Basic Books.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition, 14*, 237–273. doi:10.1016/0010-0277(83)90006-9

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

Bulman, R. J., & Wortman, C. B. (1977). Attributions of blame and coping in the "real world": Severe accident victims react to their lot. *Journal of Personality and Social Psychology, 35*, 351–363. doi:10.1037/0022-3514.35.5.351

Carey, S. (2009). *The origins of concepts*. New York, NY: Oxford University Press.

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

Carlsmith, K. M., & Sood, A. M. (2009). The fine line between interrogation and retribution. *Journal of Experimental Social Psychology, 45*(1), 191–196. doi:10.1016/j.jesp.2008.08.025

Carruthers, P. (1996). Autism as mindblindness: An elaboration and partial defence. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 257–273). Cambridge, UK: Cambridge University Press. Retrieved from http://cogprints.org/1193/0/autism.htm

Castano, E., & Giner-Sorolla, R. (2006). Not quite human: Infrahumanization in response to collective responsibility for intergroup killing. *Journal of Personality and Social Psychology, 90*, 804–818. doi:10.1037/0022-3514.90.5.804

Chalmers, D. J. (1997). *The conscious mind: In search of a fundamental theory*. New York, NY: Oxford University Press.

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences*, *108*(3), 15655–15659. doi:10.1073/pnas.1010658108

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: evidence for the oral origins of moral disgust. *Science (New York, N.Y.), 323*, 1222–1226. doi:10.1126/science.1165565

Church, R. M. (1959). Emotional reactions of rats to the pain of others. *Journal of Comparative and Physiological Psychology, 52*, 132–134.

Cialdini, R. B., Kenrick, D. T., & Hoerig, J. H. (1976). Victim derogation in the Lerner paradigm: Just world or just justification?

Journal of Personality and Social Psychology, 33, 719–724. doi:10.1037/0022-3514.33.6.719

Ciaramelli, E., Muccioli, M., Làdavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience, 2*, 84–92. doi:10.1093/scan/nsm001

Cikara, M., Eberhardt, J. L., & Fiske, S. T. (2010). From agents to objects: Sexist attitudes and neural responses to sexualized targets. *Journal of Cognitive Neuroscience*, *23*(3), 540–551. doi:10.1162/jocn.2010.21497

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

Croll, E. (2000). *Endangered daughters: Discrimination and development in Asia*. London, UK: Routledge.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*, 353–380. doi:10.1016/j.cognition.2008.03.006

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.1467-9280.2006.01834.x

Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science, 35*, 1052–1075.

Darley, J. M., & Zanna, M. P. (1982). Making moral judgments. *American Scientist, 70*, 515–521.

Davidson, D. (1970). Mental events. In L. Foster & J. W. Swanson (Eds.), *Experience and theory* (pp. 247–256). Amherst: University of Massachusetts Press.

Davis, D. A. (1991). *The Jeffrey Dahmer story: An American nightmare*. New York, NY: St. Martin's.

Davis, M. H. (1996). *Empathy: A social psychological approach*. Boulder, CO: Westview.

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion Review, 3*, 92–108. doi:10.1177/1754073910374662

Demertzi, A., Liew, C., Ledoux, D., Bruno, M.-A., Sharpe, M., Laureys, S., & Zeman, A. (2009). Dualism persists in the science of mind. *Annals of the New York Academy of Sciences, 1157*, 1–9. doi:10.1111/j.1749-6632.2008.04117.x

Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1997). *Kinds of minds: Toward an understanding of consciousness*. New York, NY: Basic Books.

DeScioli, P. (2008). *Investigations into the problems of moral cognition*. Philadelphia, PA: University of Pennsylvania. Retrieved from http://ezp-prod1.hul.harvard.edu/

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition, 112*(2), 281–299. doi:10.1016/j.cognition.2009.05.008

De Soto, C. B. (1960). Learning a social structure. *The Journal of Abnormal and Social Psychology, 60*, 417–421. doi:10.1037/h0047511

de Waal, F. B. M. (2006). *Our inner ape: The best and worst of human nature*. London, UK: Granta.

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

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 50, pp. 307–338). Retrieved from http://www.sciencedirect.com/

Eibach, R. P., Libby, L. K., & Gilovich, T. D. (2003). When change in the self is mistaken for change in the world. *Journal of Personality and Social Psychology, 84*, 917–931. doi:10.1037/0022-3514.84.5.917

Epley, N., Caruso, E., & Bazerman, M. H. (2006). When perspective taking increases taking: Reactive egoism in social interaction.

Journal of Personality and Social Psychology, 91, 872–889. doi:10.1037/0022-3514.91.5.872

Epley, N., & Waytz, A. (2010). Mind perception. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology (*5th ed., pp. 498–541). New York, NY: Wiley.

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

Esses, V. M., Veenvliet, S., Hodson, G., & Mihic, L. (2008). Justice, morality, and the dehumanization of refugees. *Social Justice Research, 21*, 4–25. doi:10.1007/s11211-007-0058-4

Fincham, F., & Jaspars, J. (1979). Attribution of responsibility to the self and other in children and adults. *Journal of Personality and Social Psychology, 37*, 1589–1602. doi:10.1037/0022-3514.37.9.1589

Fincham, F. D., & Roberts, C. (1985). Intervening causation and the mitigation of responsibility for harm doing II: The role of limited mental capacities. *Journal of Experimental Social Psychology, 21*(2), 178–194. doi:10.1016/0022-1031(85)90014-9

Finkel, N. J., & Parrott, W. G. (2006). *Emotions and culpability: How the law is at odds with psychology, jurors, and itself.* Washington, DC: American Psychological Association.

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

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy, 68*, 5–20. doi:10.2307/2024717

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*, 459–473. doi:10.1098/rstb.2002.1218

Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences, 34*, 795–817. doi:10.1016/S0191-8869(02)00072-7

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology, 78*, 708–724. doi:10.1037/0022-3514.78.4.708

Gall, T. L. (2004). Relationship with God and the quality of life of prostate cancer survivors. *Quality of Life Research, 13*, 1357–1368.

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

Gendler, T. S. (2008b). Alief in action (and reaction). *Mind & Language, 23*, 552–585. doi:10.1111/j.1468-0017.2008.00352.x

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193. doi:10.1016/0010-0277(95)00661-H

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Gilligan, C. (1993). *In a different voice: Psychological theory and women's development (*29th ed.*).* Cambridge, MA: Harvard University Press.

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

Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology, 94*, 292–306. doi:10.1037/0022-3514.94.2.292

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language, 7*(1–2), 145–171. doi:10.1111/j.1468-0017.1992.tb00202.x

Gorney, C. (2011, June). Too young to wed: The secret world of child brides. *National Geographic,* pp. 78–99.

Gould, R. V. (2003). *Collision of wills: How ambiguity about social rank breeds conflict.* Chicago, IL: University of Chicago Press.

Graham, J. (2011). *Age relations with moral foundations in a nationally-representative dataset [Unpublished raw data].* University of Southern California.

Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review, 14*, 140–150. doi:10.1177/1088868309353415

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

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366–385. doi:10.1037/a0021847

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

Gray, K. (in press). The Power of good intentions: Perceived benevolence soothes pain, increase pleasure, and improves taste. *Social Psychological and Personality Science.*

Gray, K. (2010a). *Cruel shocks and kind massages: Pain, pleasure and perceived intention* (Unpublished doctoral dissertation). Harvard University, Cambridge, MA.

Gray, K. (2010b). Moral transformation. *Social Psychological and Personality Science, 1*, 253–258. doi:10.1177/1948550610367686

Gray, K., Gray, H. M., & Wegner, D. M. (2008). *A replication of the mind survey including Google [Unpublished raw data].* Harvard University.

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

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

Gray, K., Knobe, J., Sheskin, M., Bloom, P., & Barrett, L. F. (2011). More than a body: Mind perception and objectification. *Journal of Personality and Social Psychology*, *101*(6), 1207–1220.

Gray, K., Knobe, J., Sheskin, M., Bloom, P., & Barrett, L. F. (2011). *Perceptions of mind to clothed and naked targets across cultures [Unpublished raw data].* University of Maryland, Maryland, MD.

Gray, K., & Ward, A. F. (2011). *The harm hypothesis: Perceived harm unifies morality.* Manuscript under review.

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

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

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

Gray, K., & Wegner, D. M. (2010b). Torture and judgments of guilt. *Journal of Experimental Social Psychology, 46*, 233–235. doi:10.1016/j.jesp.2009.10.003

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

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

Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., & Lyon, D. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview.

Journal of Personality and Social Psychology, 58, 308–318. doi:10.1037/0022-3514.58.2.308

Greene, J. D. (2007). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 3: The neuroscience of morality: emotion, disease, and development* (pp. 35–80). Cambridge, MA: MIT Press.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*, 389–400.

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*, 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.

Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective science* (pp. 852–870). Oxford, UK: Oxford University Press.

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

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*(1), 98–116. doi:10.1007/s11211-007-0034-z

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

Haidt, J., Rozin, P., Mccauley, C., & Imada, S. (1997). Body, psyche, and culture: The relationship between disgust and morality. *Psychology & Developing Societies, 9*, 107–131. doi:10.1177/097133369700900105

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557–559. doi:10.1038/nature06288

Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*, 843–855. doi:10.1111/j.1467-8624.1995.tb00909.x

Hare, R. D. (1998). *Without conscience: The disturbing world of the psychopaths among us*. New York, NY: Guilford.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low. *Psychological Science, 17*, 847–853. doi:10.1111/j.1467-9280.2006.01793.x

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

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

Hauser, M. (2006). *Moral minds: How nature designed a universal sense right and wrong*. New York, NY: HarperCollins.

Heflick, N. A., & Goldenberg, J. L. (2009). Objectifying Sarah Palin: Evidence that objectification causes women to be perceived as less competent and less fully human. *Journal of Experimental Social Psychology, 45*, 598–601. doi:10.1016/j.jesp.2009.02.008

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360–1366. doi:10.1126/science.1146282

Huebner, B., Bruno, M., & Sarkissian, H. (2009). What does the nation of china think about phenomenal states? *Review of Philosophy and Psychology, 1*, 225–243. doi:10.1007/s13164-009-0009-0

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Oxford, UK: Cambridge University Press.

Humphrey, N. (2003). *The mind made flesh: Essays from the frontiers of psychology and evolution*. New York, NY: Oxford University Press.

Husseini, R. (2009). *Murder in the name of honor*. Oxford, UK: Oneworld.

Janoff-Bulman, R. (1979). Characterological versus behavioral self-blame: Inquiries into depression and rape. *Journal of Personality and Social Psychology, 37*, 1798–1809. doi:10.1037/0022-3514.37.10.1798

Janoff-Bulman, R. (2009). To provide or protect: Motivational bases of political liberalism and conservatism. *Psychological Inquiry, 20*, 120. doi:10.1080/10478400903028581

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: Two faces of moral regulation. *Journal of Personality and Social Psychology, 96*, 521–537. doi:10.1037/a0013779

Janoff-Bulman, R., Timko, C., & Carli, L. L. (1985). Cognitive biases in blaming the victim. *Journal of Experimental Social Psychology, 21*, 161–177. doi:16/0022–1031(85)90013–7

Jenkins, A. C., & Wegner, D. M. (2011). *Modes of mind perception*. Manuscript in preparation.

Johnson-Laird, P. N. (1986). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Jones, E. E., & Davis, K. E. (1966). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York, NY: Academic Press. Retrieved from http://www.sciencedirect.com/

Jost, J. T., Napier, J. L., Thorisdottir, H., Gosling, S. D., Palfai, T. P., & Ostafin, B. (2007). Are needs to manage uncertainty and threat associated with political conservatism or ideological extremity? *Personality and Social Psychology Bulletin, 33*, 989–1007. doi:10.1177/0146167207301028

Kahneman, D., Schkade, D., & Sunstein, C. (1998). Shared outrage and erratic awards: The psychology of punitive damages. *Journal of Risk and Uncertainty, 16*(1), 49–86. doi:10.1023/A:1007710408413

Karniol, R. (1978). Children's use of intention cues in evaluating behavior. *Psychological Bulletin, 85*, 76–85. doi:10.1037/0033-2909.85.1.76

Kay, A. C., Gaucher, D., McGregor, I., & Nash, K. (2010). Religious belief as compensatory control. *Personality and Social Psychology Review, 14*, 37–48. doi:10.1177/1088868309353750

Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition & Emotion, 10*, 155–171. doi:10.1080/026999396380312

Kervyn, N., Yzerbyt, V. Y., Demoulin, S., & Judd, C. M. (2008). Competence and warmth in context: The compensatory nature of stereotypic views of national groups. *European Journal of Social Psychology, 38*, 1175–1183. doi:10.1002/ejsp.526

Killen, M., Lynn Mulvey, K., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition, 119*, 197–215. doi:10.1016/j.cognition.2011.01.006

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis, 63*, 190–193.

Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences, 7*, 67–83. doi:10.1007/s11097-007-9066-y

Knobe, J., Jenkins, A. C., Dodell-Feder, D., & Saxe, R. (2011). *The group has a mind of itsown: Common neural bases for individual and group theory-of-mind*. Manuscript under review.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature, 446*, 908–911. doi:10.1038/nature05631

Kohlberg, L. (1981). *The philosophy of moral development: Moral stages and the idea of justice*. New York, NY: HarperCollins.

Kolenda, P. M. (1964). Religious anxiety and Hindu fate. *The Journal of Asian Studies, 23*, 71–81. doi:10.2307/2050622

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology, 90*, 543–555. doi:10.1037/0022-3514.90.4.543

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480–498.

Lane, J. D., Wellman, H. M., & Evans, E. M. (2010). Children's understanding of ordinary and extraordinary minds. *Child Development, 81*, 1475–1489. doi:10.1111/j.1467-8624.2010.01486.x

Lerner, M. (1980). *The belief in a just world: A fundamental delusion*. New York, NY: Springer.

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin, 85*, 1030–1051. doi:10.1037/0033-2909.85.5.1030

Lerner, M. J., & Simmons, C. H. (1966). Observer's reaction to the "innocent victim": Compassion or rejection? *Journal of Personality and Social Psychology, 4*, 203–210. doi:10.1037/h0023562

Lewis, G. (1995). The articulation of circumstance and causal understanding. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 557–574). New York, NY: Oxford University Press.

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review, 4*, 186–197. doi:10.1207/S15327957PSPR0402_06

Lombrozo, T. (2010). Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology, 61*, 303–332.

Loughnan, S., Haslam, N., & Bastian, B. (2010). The role of meat consumption in the denial of moral status and mind to meat animals. *Appetite, 55*(1), 156–159. doi:10.1016/j.appet.2010.05.043

Loughnan, S., Haslam, N., Murnane, T., Vaes, J., Reynolds, C., & Suitner, C. (2010). Objectification leads to depersonalization: The denial of mind and moral concern to objectified others. *European Journal of Social Psychology, 40*, 709–717. doi:10.1002/ejsp.755

Mahmut, M. K., Homewood, J., & Stevenson, R. J. (2008). The characteristics of non-criminals with high psychopathy traits: Are they similar to criminal psychopaths? *Journal of Research in Personality, 42*, 679–692. doi:10.1016/j.jrp.2007.09.002

Malle, B. F. (2005). Folk theory of mind: Conceptual foundations of human social cognition. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 225–255). New York, NY: Oxford University Press.

Malle, B. F. (2006). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.

Malle, B. (2011). Time to give up the dogmas of attribution: An alternative theory of behavior explanation. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 44, pp. 297–352). Burlington, VT: Academic Press.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (in press). Moral, cognitive, and social: The nature of blame. In J. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social thinking and interpersonal behavior*. Philadelphia, PA: Psychology Press.

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews, 32*, 454–465. doi:10.1016/j.neubiorev.2007.08.003

Martin, G. B., & Clark, R. D. (1982). Distress crying in neonates: Species and peer specificity. *Developmental Psychology, 18*, 3–9. doi:10.1037/0012-1649.18.1.3

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19*, 242–279. doi:16/0010-0285(87)90012-0

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115. doi:10.1146/annurev.ps.32.020181.000513

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences, 11*, 143–152. doi:10.1016/j.tics.2006.12.007

Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*, 76–93.

Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., & Gabrieli, J. D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences, 108*, 2688–2692. doi:10.1073/pnas.1011734108

Morewedge, C. K. (2009). Negativity bias in attribution of external agency. *Journal of Experimental Psychology. General, 138*, 535–545. doi:10.1037/a0016796

Muentener, P., & Lakusta, L. (2011). The intention-to-CAUSE bias: Evidence from children's causal language. *Cognition, 119*, 341–355. doi:10.1016/j.cognition.2011.01.017

Mumford, S. R. (1989). *Himalayan dialogue: Tibetan lamas and Gurung shamans in Nepal*. Madison, WI: University of Wisconsin Press.

Murphy, G. L. (2004). *The big book of concepts*. Cambridge, MA: MIT Press.

Nichols, S. (2002). Norms with feeling: towards a psychological account of moral judgment. *Cognition, 84*, 221–236.

Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs, 41*, 663–685. doi:10.1111/j.1468-0068.2007.00666.x

Norenzayan, A., & Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science, 322*, 58–62. doi:10.1126/science.1158757

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*, 1560–1563. doi:10.1126/science.1133755

Nucci, L. P., Turiel, E., & Encarnacion-Gawrych, G. (1983). Children's social interactions and social concepts. *Journal of Cross-Cultural Psychology, 14*, 469–487. doi:10.1177/0022002183014004006

Nussbaum, M. C. (1995). Objectification. *Philosophy and Public Affairs, 24*, 249–291.

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, NY: Routledge.

Osofsky, M. J., Bandura, A., & Zimbardo, P. G. (2005). The role of moral disengagement in the execution process. *Law and Human Behavior, 29*, 371–393. doi:10.1007/s10979-005-4930-1

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.

Pargament, K. I., Ensing, D. S., Falgout, K., Olsen, H., Reilly, B., Haitsma, K., & Warren, R. (1990). God help (I): Religious coping efforts as predictors of the outcomes to significant negative life events. *American Journal of Community Psychology, 18*, 793–824. doi:10.1007/BF00938065

Parkinson, C., Sinnott-Armstrong, W., Koralus, P. E., Mendelovici, A., McGeer, V., & Wheatley, T. (2011). Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust. *Journal of Cognitive Neuroscience*, *23*(10), 3162–3180. doi:10.1162/jocn_a_00017

Patterson, C. (1985). "Not worth the rearing": The causes of infant exposure in ancient greece. *Transactions of the American Philological Association, 115*, 103–123. doi:10.2307/284192

Paulhus, D. L., Hemphill, J., & Hare, R. (2009). *Manual of the Self-Report Psychopathy Scale (SRP–III)*. Toronto, Canada: Multi-Health Systems.

Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology, 95*, 274–281. doi:10.1037/0021-843X.95.3.274

Pepitone, A., & Saffiotti, L. (1997). The selectivity of nonmaterial beliefs in interpreting life events. *European Journal of Social Psychology, 27*(1), 23–35. doi:10.1002/(SICI)1099-0992(199701)27:1<23::AID-EJSP805>3.0.CO;2-B

Phillips, J., & Knobe, J. (2009). Moral judgments and intuitions about freedom. *Psychological Inquiry, 20*, 30–36.

Piaget, J. (1932). *The moral judgment of the child*. New York, NY: Harcourt Brace.

Pinker, S. (2011). *The better angels of our nature: Why violence has declined.* New York, NY: Viking.

Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 91–108). Washington, DC: APA Press.

Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science, 14*, 267–272. doi:10.1111/1467-9280.03433

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515–526. doi:10.1017/S0140525×00076512

Preston, S. D., & de Waal, F. B. M. (2001). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1–20. doi:10.1017/S0140525×02000018

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

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Reeder, G. D. (2009). Mindreading: Judgments about intentionality and motives in dispositional inference. *Psychological Inquiry, 20*, 1–18. doi:10.1080/10478400802615744

Reverby, E.-S. M. (2000). *Tuskegee's truths: Rethinking the Tuskegee Syphilis Study*. Chapel Hill, NC: The University of North Carolina Press.

Ridley, M. (1998). *The Origins of Virtue: Human Instincts and the Evolution of Cooperation* (1st ed.). London, UK: Penguin.

Robbins, P., & Jack, A. I. (2006). The phenomenal stance. *Philosophical Studies, 127*(1), 59–85. doi:10.1007/s11098-005-1730-x

Rochat, P., Striano, T., & Morgan, R. (2004). Who is doing what to whom? Young infants' developing sense of social causality in animated displays. *Perception, 33*, 355–369. doi:10.1068/p3389

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin, 121*, 133–148. doi:10.1037/0033-2909.121.1.133

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605. doi:10.1016/0010-0285(75)90024-9

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989). Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology, 57*, 681–690. doi:10.1037/0022-3514.57.4.681

Routledge, C., & Arndt, J. (2008). Self sacrifice as self defence: Mortality salience increases efforts to affirm a symbolic immortal self at the expense of the physical self. *European Journal of Social Psychology, 38*, 531–541. doi:10.1002/ejsp.442

Royzman, E. B., Leeman, R. F., & Baron, J. (2009). Unsentimental ethics: Towards a content-specific account of the moral-conventional distinction. *Cognition, 112*(1), 159–174. doi:10.1016/j.cognition.2009.04.004

Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust. In *Handbook of emotions* (3rd ed., pp. 757–776). New York, NY: Guilford.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178. doi:10.1037/h0077714

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology, 12*, 175–176. doi:10.1037/0012-1649.12.2.175

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16*, 235–239. doi:10.1016/j.conb.2006.03.001

Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science, 16*, 995–1001. doi:10.1111/j.1467-9280.2005.01649.x

Schmidt, G., & Weiner, B. (1988). An attribution-affect-action theory of behavior: Replications of judgments of help-giving. *Personality and Social Psychology Bulletin, 14*, 610–621. doi:10.1177/0146167288143021

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

Sen, M. (2002). *Death by fire: Sati dowry death and infanticide in modern India*. London, UK: Phoenix.

Shariff, A. F., & Norenzayan, A. (2007). God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Science, 18*, 803–809. doi:10.1111/j.1467-9280.2007.01983.x

Shaver, K. G. (1985). *The attribution of blame: Causality, reponsibility, and blameworthiness*. New York, NY: Springer-Verlag.

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: when cheating leads to moral disengagement and motivated forgetting. *Personality & Social Psychology Bulletin, 37*, 330–349. doi:10.1177/0146167211398138

Shweder, R. A., Mahapatra, M., & Miller, J. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 1–83). Chicago, IL: University of Chicago Press.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, and divinity), and the "big three" explanations of suffering. In *Morality and health* (pp. 119–169). New York, NY: Routledge.

Singer, P. (1975). *Animal liberation*. New York, NY: Random House.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303*, 1157–1162. doi:10.1126/science.1093535

Sinnott-Armstrong, W., & Wheatley, T. (2011). *Moral judgments are not unified: Why moral psychology needs a taxonomic approach*. Manuscript under review.

Skitka, L. J., & Mullen, E. (2002). Understanding judgments of fairness in a real-world political context: A test of the value protection model of justice reasoning. *Personality and Social Psychology Bulletin, 28*, 1419–1429. doi:10.1177/014616702236873

Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty, 26*, 5–16.

Smith, A. (2009). The empathy imbalance hypothesis of Autism: A theoretical approach to cognitive and emotional empathy

123

in autistic development. *The Psychological Record, 59*, 489–510.

Spilka, B., & Schmidt, G. (1983). General attribution theory for the psychology of religion: The influence of event-character on attributions to god. *Journal for the Scientific Study of Religion, 22*, 326–339. doi:10.2307/1385771

Spilka, B., Zwartjes, W. J., & Zwartjes, G. M. (1991). The role of religion in coping with childhood cancer. *Pastoral Psychology, 39*, 295–304. doi:10.1007/BF01020701

Sripada, C. S. (2009). The Deep Self Model and asymmetries in folk judgments about intentional action. *Philosophical Studies, 151*, 159–176. doi:10.1007/s11098-009-9423-5

Stotland, E. (1969). Exploratory investigations of empathy. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 4, pp. 271–314). New York, NY: Academic Press.

Strohminger, N., Lewis, R. L., & Meyer, D. E. (2011). Divergent effects of different positive emotions on moral judgment. *Cognition, 119*, 295–300. doi:16/j.cognition.2010.12.012

Suhler, C. L., & Churchland, P. (2011). Can innate, modular "foundations" explain morality? Challenges for Haidt's moral foundations theory. *Journal of Cognitive Neuroscience,* 1–14. doi:10.1162/jocn.2011.21637

Sytsma, J., & Machery, E. (2009). Two conceptions of subjective experience. *Philosophical Studies, 151*, 299–327. doi:10.1007/s11098-009-9439-x

Tannenbaum, D., Uhlmann, E. L., & Diermeier, D. (in press). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology*.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin, 110*, 67–85.

Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology, 52*, 700–709. doi:10.1037/0022-3514.52.4.700

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–691. doi:10.1017/S0140525×05000129

Tsoulis-Reay, A. (2010, December 10). Is incest a two-way street? David Epstein is charged with having sex with his adult daughter. Isn't she guilty, too? *Slate*. Retrieved from http://www.slate.com/id/2277681/

Tulloch, M. (2000). The meaning of age: Differences in the fear of crime. *British Journal of Criminology, 40*, 451–467. doi:10.1093/bjc/40.3.451

Turiel, E. (1983). *The development of social knowledge: morality and convention*. Cambridge, UK: Cambridge University Press.

Turiel, E., Hildebrandt, C., Wainryb, C., & Saltzstein, H. D. (1991). Judging social issues: Difficulties, inconsistencies, and consistencies. *Monographs of the Society for Research in Child Development, 56*, 1–116. doi:10.2307/1166056

Valdesolo, P., & DeSteno, D. (2011). The virtue in vice: Short-sightedness in the study of moral emotions. *Emotion Review, 3*, 276–277. doi:10.1177/1754073911402390

van Leeuwen, F., & Park, J. H. (2011). *Moral virtues and political orientation: Assessing priorities under constraint*. Unpublished manuscript, University of Bristol, Bristol, UK.

Van Overwalle, F., Van Duynslaeger, M., Coomans, D., & Timmermans, B. (2012). Spontaneous goal inferences are often inferred faster than spontaneous trait inferences. *Journal of Experimental Social Psychology, 48*(1), 13–18.

Walker, L. J., Frimer, J. A., & Dunlop, W. L. (2010). Varieties of moral personality: Beyond the banality of heroism. *Journal of Personality, 78*, 907–942. doi:10.1111/j.1467-6494.2010.00637.x

Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology, 3*, 73–79. doi:10.1037/h0022733

Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who sees human? *Perspectives on Psychological Science, 5*, 219–232. doi:10.1177/1745691610369336

Waytz, A., & Young, L. (2012). The group-member mind tradeoff: Attributing mind togroups versus group members. *Psychological Science*, *23*, 77–85.

Wegner, D. M., & Vallacher, R. R. (1977). *Implicit psychology: An introduction to social cognition*. New York, NY: Oxford University Press.

Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology, 39*, 186–200.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of conduct of social conduct*. New York, NY: Guilford Press.

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

Wiggins, J. S., & Broughton, R. (1991). A geometric taxonomy of personality scales. *European Journal of Personality, 5*, 343–365.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34. doi:10.1016/S0010-0277(98)00058-4

Woolfolk, R. L., Doris, J. M., & Darley, J. M. (2006). Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility. *Cognition, 100*, 283–301. doi:16/j.cognition.2005.05.002

Wright, J. C., & Baril, G. (in press). The role of cognitive resources in determining our moral intuitions: Are we all liberals at heart? *Journal of Experimental Social Psychology*. doi:10.1016/j.jesp.2011.03.014

Ybarra, O. (2002). Naive causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin, 128*, 421–441. doi:10.1037/0033-2909.128.3.421

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences, 107*, 6753–6758. doi:10.1073/pnas.0914826107

Young, L., & Phillips, J. (2011). The paradox of moral focus. *Cognition, 119*, 166–178. doi:10.1016/j.cognition.2011.01.004

Young, L., & Saxe, R. (2011). The role of intent for distinct moral domains. *Cognition*, *120*, 202–214.

Yuill, N. (1984). Young children's coordination of motive and outcome in judgements of satisfaction and morality. *British Journal of Developmental Psychology, 2*(1), 73–81.

Zalla, T., Machery, E., & Leboyer, M. (2008, January). *Intentional action and moral judgmentin Asperger Syndrome and high-functioning Autism*. Paper presented at the Winter Workshop 2008 on Games, Experiments and Philosophy, Jena, Germany.

Zigler, E. (1967). Familial mental retardation: A continuing dilemma. *Science, 155*, 292–298. doi:10.1126/science.155.3760.292

Zimbardo, P. (2008). *The Lucifer Effect: Understanding how good people turn evil*. New York, NY: Random House.

Zitek, E. M., Jordan, A. H., Monin, B., & Leach, F. R. (2010). Victim entitlement to behave selfishly. *Journal of Personality and Social Psychology, 98*, 245–255. doi:10.1037/a0017168