The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm

Chelsea Schein[1] and Kurt Gray[1]

Abstract

The nature of harm—and therefore moral judgment—may be misunderstood. Rather than an objective matter of reason, we argue that harm should be redefined as an intuitively perceived continuum. This redefinition provides a new understanding of moral content and mechanism—the constructionist Theory of Dyadic Morality (TDM). TDM suggests that acts are condemned proportional to three elements: norm violations, negative affect, and—importantly—perceived harm. This harm is dyadic, involving an intentional agent causing damage to a vulnerable patient (A→P). TDM predicts causal links both from harm to immorality (dyadic comparison) and from immorality to harm (dyadic completion). Together, these two processes make the "dyadic loop," explaining moral acquisition and polarization. TDM argues against intuitive harmless wrongs and modular "foundations," but embraces moral pluralism through varieties of values and the flexibility of perceived harm. Dyadic morality impacts understandings of moral character, moral emotion, and political/cultural differences, and provides research guidelines for moral psychology.

> *The man intentionally gished the little girl, who cried.*

Gish is not a real word, but when people read this sentence, they automatically judge the man to be doing something immoral.[1] What causes this intuitive moral judgment? Theories of morality often invoke the intrinsic wrongness of certain acts, but here the act is unclear. Instead, people are reacting to the structure of the act: an intentional agent is acting upon a vulnerable victim—and ostensibly causing her harm (Hester, Payne, & Gray, 2017). Perceptions of harm not only cause moral judgment here, but may also underlie moral judgment in general.

Harm drives the moral condemnation of murder, assault, rape, and theft (Haidt, Koller, & Dias, 1993; Mikhail, 2007), and—if moral rhetoric is taken at face value—harm also drives the condemnation of pornography (Pierce, 2001), trigger warnings (Lukianoff & Haidt, 2015), homosexuality (Bryant, 1977), vaccinations (Mnookin, 2012), transgender bathroom use (Harrison, 2016), dishonor (Nisbett & Cohen, 1996), and even reading Harry Potter (James, 2010). In his guide to human health, Dr. Kellogg condemned masturbation based on the "scientific" claim that "neither the plague, nor war, nor small-pox, nor similar diseases, have produced results so disastrous to humanity as the pernicious habit of onanism" (Dr. Adam Clarke on masturbation, quoted in Kellogg, 1890, p. 233). The creator of Kellogg's cereal was claiming that masturbation was morally wrong because of the harm it caused.

Perceptions of harm are ubiquitous in moral judgments and there are two possibilities for why. The first possibility—currently favored in the field—is that perceptions of harm are epiphenomenal spasms of post hoc reasoning, reflecting only the moral mind's attempt to justify itself (Haidt, 2001, 2012). The second possibility—explored here—is that harm is the fundamental basis by which moral judgments are made and maintained.

Précis of Dyadic Morality

The Theory of Dyadic Morality (TDM) suggests that moral judgments revolve around a cognitive template of harm. In dyadic morality—and throughout this article—harm has a very specific definition: It involves two perceived and causally connected minds, *an intentional agent causing damage*

[1]The University of North Carolina at Chapel Hill, USA

Corresponding Author:
Kurt Gray, Department of Psychology and Neuroscience, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
Email: kurtgray@unc.edu

Figure 1. Moral judgment revolves around a fuzzy cognitive template of harm, defined as the perception of an intentional agent causing damage to a vulnerable patient.

to a vulnerable patient (see Figure 1), denoted as $\left(iA \overset{d}{\to} vP \right)$,

or more simply (A→P). As we explore, this harm-based cognitive template functions intuitively and is rooted in innate and evolved processes of the human mind; it is also shaped by cultural learning, therefore allowing cultural pluralism. TDM therefore reconciles social intuitionism (Haidt, 2001) with developmental harm-centric theories (Turiel, 1983), while also incorporating insights from modern pluralistic accounts (Haidt, 2012). Dyadic morality is also compatible with constructionist accounts of emotion and cognition (Lindquist, 2013) and with modern domain-general understandings of moral and nonmoral cognition (Helion & Pizarro, 2015; Kruglanski & Gigerenzer, 2011; Miller & Cushman, in press; Van Bavel, FeldmanHall, & Mende-Siedlecki, 2015).

From Harm to Immorality, From Immorality to Harm

Dyadic morality is a dynamic causal model, suggesting bidirectional—and mutually reinforcing—causal links between harm (A→P) and moral judgment. The link from harm to immorality means that perceived harm causes acts to be judged as immoral; when acts seem harmful, they seem morally wrong.

This harm to immorality link allows for the acquisition of morality through both firsthand perception of an act's harmfulness (Blair, 1995; Smetana, 1985) and through another's testimony of harm (Rottman, Young, & Kelemen, 2017), such as a pastor detailing the suffering caused by gay marriage (Comer, 2012).
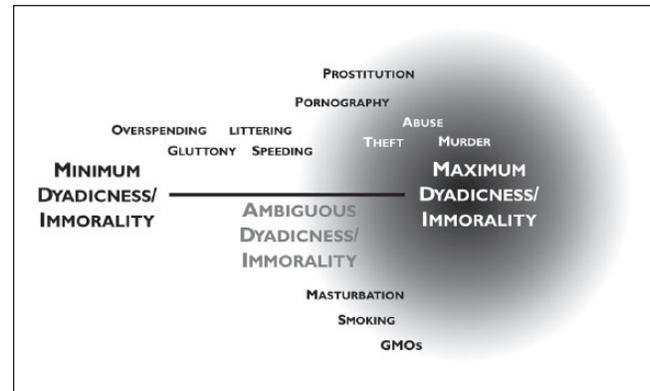

Figure 2. Dyadic morality predicts that the continuum of moral condemnation is predicted by a continuum of "dyadicness" (i.e., how dyadic an act intuitively seems to be).
Note. GMO = genetically modified food.

The complementary link from immorality to harm means that judgments of immorality also cause perceptions of harm; when acts seem morally wrong, they seem harmful. This causal link allows for moral acquisition through direct testimony of an act's wrongness (e.g., parents say an act is just wrong; Rottman et al., 2017), and entrenchment of these moral judgments through subsequent perceptions of harm.

Together, these dynamic connections between harm and immorality create a feedback cycle in which acts become iteratively moralized and "harmified" through time and social discussion (the "dyadic loop"), helping to explain moral polarization and the concurrent "creep" of harm and morality (Haslam, 2016; Schein & Gray, 2016; in press).

Causal links between harm and immorality provide the most basic prediction of dyadic morality: that the immorality of an act is fundamentally predicted by the amount of harm it involves (or implies).

Immorality clearly exists on a continuum—from overspending, to speeding, to theft, to murder—and TDM predicts that this continuum of immorality is grounded in a continuum of perceived harm or "dyadicness" (Figure 2). The more an act seems to involve an intentional agent causing damage to a vulnerable patient, the more immoral it should seem.

Importantly, this continuum of harm is a matter of intuitive perception—just like moral judgment—and not of reasoned objective fact. While reasoning can influence perceptions of harm by arguing for (or against) its existence, TDM suggest that the intuitive perception of harm is ultimately what drives moral judgment.

Harmless Wrongs?

The causal power of harm is largely uncontroversial, and the field agrees that it accounts for the most typical and universal moral judgments (Haidt, Graham, & Ditto, 2015; Hofmann, Wisneski, Brandt, & Skitka, 2014). However,

many argue for a set of "objectively harmless" moral wrongs (e.g., consensual incest) in which perceptions of harm reflect only post hoc reasoning (Haidt, 2001; Scott, Inbar, & Rozin, 2016).

One problem with these claims is that "harmless" wrongs are typically assessed with weird cases of dog-eating (Haidt et al., 1993), corpse-sexing (Haidt & Hersh, 2001), animal pornography (Eibach, Libby, & Ehrlinger, 2009), or masturbation while cuddling a teddy bear (Helzer & Pizarro, 2011)—and weirdness is tied to immorality (it represents a norm violation; Gray & Keeney, 2015b).

The major problem with research on "harmless" wrongs is that it ignores the perceptions of participants, relying instead upon the assumptions of researchers—and participants see substantial harm in "harmless" wrongs (Royzman, Kim, & Leeman, 2015). Past research also fails to assess harm and morality intuitively, employing explicit self-perception questionnaires answered with ample time (Graham & Haidt, 2012; Graham, Haidt, & Nosek, 2009).

When harm and morality are assessed intuitively, the presence of harm in "harmless wrongs" is clear (Gray, Schein, & Ward, 2014). We therefore suggest that it is judgments of harmlessness—and not harm—that require effortful reasoning, because harm and morality are naturally fused together in the human mind. Moreover, it is the presence of perceived harm that distinguishes immoral acts from those that are merely bad (Schein & Gray, 2015) or disgusting (Schein, Ritter, & Gray, 2016).

The intuitive presence of harm in "harmless" wrongs is important because these wrongs have been used to argue for moral mechanisms independent of harm—such as the "purity" module of Moral Foundations Theory (MFT; Graham et al., 2013). Dyadic morality explicitly denies the existence of these harm-independent mechanisms, a denial consistent with ample recent research (Cameron, Lindquist & Gray, 2015; Schein & Gray, 2015; Schein, Ritter, & Gray, 2016). Rather than distinct causal mechanisms that *explain* moral judgment, TDM suggests that the moral "foundations" are taxonomic categories that merely *name* different varieties of values, just as music "genres" name different varieties of music. In addition to MFT, there are multiple taxonomies that descriptively catalog political differences (Rai & Fiske, 2011; Schwartz & Bilsky, 1987), but these researcher-compiled inventories should not be confused with fundamental psychological processes.

Dyadic Pluralism

Dyadic morality embraces moral pluralism. TDM acknowledges that cultures moralize different values (Rachels, 1986; Shweder, 2012), but predicts that moral pluralism is underlain by varying perceptions of harm (e.g., conservatives moralize patriotism and chastity because they view their violation as harmful—unlike liberals). TDM therefore supports diverse moral concerns such as loyalty, purity, industriousness, and social order, but suggests that they are best understood as "transformations" or "intermediaries" of harm, values whose violation leads to perceptions of concrete harm. For example, Anita Bryant believed that the "purity" violation of gay rights would convince kids to be gay, which would not only destroy their vulnerable immortal souls but also undermine procreation and hence the American family, which would bankrupt the nation and eventually lead to anarchy (Bryant, 1977).

More succinctly, moral pluralism is enabled through harm pluralism, which in turn is enabled by the fact that harm is synthetic—constructed from the three elements of an agent, a patient, and their causal link. Each of these elements can take on different forms depending on the culture and context: Agents can be anything perceived to have intentions such as a person, a corporation (Rai & Diermeier, 2015), a God (Gray & Wegner, 2010a), or a conspiratorial governmental agency (Epperson, 1985); patients can be anything perceived to have vulnerability (Bentham, 1970/1780; Crimston, Bain, Hornsey, & Bastian, 2016; Singer, 1981) such as children (Dijker, 2010), adults, people in a group (Cooley et al, 2017), future generations, animals (Singer, 1975), or the environment (Bastian & Crimston, 2016); and the specific act can cause damage through physical destruction, mental suffering, or spiritual defilement. Importantly, as people disagree about the potential sources or recipients of harm, TDM helps to explain moral disagreement.

Dyadic morality therefore suggests that morality is about harm, but not "just" harm. Consider origami, in which infinitely different shapes emerge by folding the same simple piece of paper; origami is paper but it's not "just" paper. With morality, infinitely different values—not just four, five, or six—may be moralized through the combinatorial flexibility of perceived harm. Dyadic morality further suggests that morality is more than "just" harm by acknowledging the importance of both norms (Monroe, Guglielmo, & Malle, 2012; Nichols, 2002; Sripada & Stich, 2005) and affect (Greene, 2013; Haidt, 2003; Hutcherson & Gross, 2011), two powerful elements in moral judgment.

Despite the diversity of perceived harm, some acts involve relatively obvious harm because they represent proximate evolutionary threats that directly endanger survival and procreation. This explains why people share relatively universal moral intuitions about murder and abuse (Mikhail, 2009). Conversely, some acts involve less obvious harm because they represent distal evolutionary threats—they endanger survival and procreation indirectly by threatening cultural coordination and social status. This ambiguity of harm explains the substantial moral disagreement engendered by prostitution, religious adherence, patriotism, censorship, immigration, and taxation (e.g., Dworkin, 1981; vs. Miller-Young, 2013).

Grounding moral judgment in perceived harm provides a new—and hopeful—perspective on moral disagreement,

suggesting that liberals and conservatives share fundamentally the same moral mind, but just see harm differently (Landy, 2016; Schein & Gray, 2015). By helping people understand that "the other side" respects the sanctity of harm, dyadic morality may help reduce the vindictiveness of moral conversations. Perhaps moral chasms can be crossed with bridges of perceived harm.

Outline

This article presents an in-depth review of dyadic morality. Dyadic morality has been introduced before (Gray, Waytz, & Young, 2012; Gray & Wegner, 2011b; Gray, Young, & Waytz, 2012), but at that time, its claims outpaced the evidence—a point noted by many commentators (Alicke, 2012; Ditto, Liu, & Wojcik, 2012; Haslam, 2012; Monroe et al., 2012). Recent empirical work (Cheng, Ottati, & Price, 2013; Clark, Chen, & Ditto, 2015; Gray, Schein, & Ward, 2014; Royzman et al., 2015; Schein & Gray, 2015; Schein, Ritter, & Gray, 2016) provide the opportunity to narrow the gap between data and theory. This article provides an updated version of TDM, one that specifies key constructs, integrates related literature, corrects misconception, and lays out testable future predictions.

The article is structured into two main sections. The first covers *moral content*—reviewing five key questions (and their answers): (a) *What makes an Act Immoral?* (Norms, Affect, and Harm), (b) *What is Harm* (Synthetic, Perceived, Intuitive, a Continuum), (c) *What is the Moral Dyad?* (A Fuzzy Cognitive Template), (d) *Are there Harmless Wrongs?* (No: Speaking for Moral Dumbfounding), and (e) *What about other Moral Values?* (They are Transformations and Intermediaries of Harm).

The second section covers *moral mechanism*—reviewing five more key questions (and their answers): (a) *How do we Make Moral Judgments?* (Through Dyadic Comparison and Direct Learning), (b) *How do Moral Judgments Impact Perceptions?* (Through Dyadic Completion), (c) *How are Moral Judgments Extended and Entrenched?* (Through The Dyadic Loop), (d) *How does Dyadic Morality Relate to MFT?* (It Argues Against Its Claim of Modularity), and (e) *Is There an Alternative to Modularity?* (Yes: Constructionism).

After the second section, we briefly consider implications of dyadic morality and how TDM might be further tested.

Moral Content

Philosophy is well known for asking definitional questions, and as moral psychology grew out of philosophy, one frequent question is, "What is morality?" often formulated as "What content separates the (im)moral from the non-moral?" (Gert, 2015; Machery, 2012). In exploring questions about moral content, we also explore questions of harm, including how best to understand this idea, whether there are harmless wrongs, and how harm interacts with various moral values.

What Makes an Act Immoral? Norms, Affect, and Harm

Defining Immorality

Although any complex concept defies a complete philosophical definition (Gert, 2015; Wittgenstein, 1953), one popular working definition of moral judgment is "evaluations (good vs. bad) of the actions or character of a person that are made with respect to a set of virtues held to be obligatory by a culture or subculture" (Haidt, 2001, p. 817). Empirical research reveals that moral judgments often have unique signatures: Compared with nonmoral judgments, moral judgments are held with deeper conviction (Skitka, Bauman, & Sargis, 2005; Skitka, Washburn, & Carsel, 2015), are often less amenable to compromise and trade-offs (Ginges, Atran, Medin, & Shikaki, 2007; Tetlock, Kristel, Beth, Green, & Lerner, 2000), are more motivating and legitimizing of behavior (Effron & Miller, 2012; Ryan, 2014; Skitka & Wisneski, 2011; van Zomeren, Postmes, & Spears, 2012), are more likely to bind together social groups (Graham & Haidt, 2010; Haidt, 2012, Chapter 9; Janoff-Bulman & Carnes, 2013), are seen as more objective and universal (Goodwin & Darley, 2008, 2012; Nichols & Folds-Bennett, 2003; Wainryb, Shaw, Langley, Cottam, & Lewis, 2004), and more independent of authority (Nucci & Turiel, 1993; Turiel, Killen, & Helwig, 1987; Wisneski, Lytle, & Skitka, 2009; though see Piazza, Sousa, & Holbrook, 2013). Immoral acts are also seen as intrinsically deserving of blame (Malle, Guglielmo, & Monroe, 2014) and punishment (Hart, 2008; Shultz, Schleifer, & Altman, 1981; Skitka & Houston, 2001), and as intrinsically tied to outrage and other negative emotions.

We suggest that an analogy for moral judgments is fire: affectively hot (Gray & Wegner, 2011a; Haidt, 2003; Hutcherson & Gross, 2011), igniting strong views (Skitka et al., 2005), and even serving as a common focus for ritual that can both bring groups together (Graham & Haidt, 2010; Greene, 2013) or be used to destroy (Atran & Ginges, 2012; Halevy, Kreps, Weisel, & Goldenberg, 2015). Fire requires three components to ignite—heat, oxygen, and fuel; moral judgment may also require three components.

Integrating diverse research (e.g., Gray, Young, & Waytz, 2012; Haidt, 2001; Nichols, 2002; Turiel et al., 1987), we suggest that moral judgments emerge when three components are present: norm violations, negative affect, and dyadic harm. These components are not as distinct as those of fire, but are causally interconnected. Norm violations are negative and may prompt perceptions of harm, negative affect predisposes people to see norm violations and harm, and dyadic harm is typically a norm violation and leads to negative affect. All these components seem to be present in moral judgments.

Norms

Norms are expectancies (Bicchieri & Muldoon, 2014; Schelling, 1981), beliefs (Cialdini, Kallgren, & Reno, 1991),

values (Nichols, 2002; Parsons, 1951), and rules (Sripada & Stich, 2005) about how other people act (descriptive) or should act (injunctive; Reno, Cialdini, & Kallgren, 1993). As such, moral violations always violate norms (Malle et al., 2014; Monroe et al., 2012; Nichols, 2002). These norms can be prescriptive (what people should do) and proscriptive (what people should not do; Janoff-Bulman, Sheikh, & Hepp, 2009), with the latter most frequently tied to behaviors (Kochanska, Coy, & Murray, 2001) and to moral judgment: the most memorable of the ten commandments govern what though "shalt not" do. Some definitions of norms involve the consequences for breaking them—which include punishment (Axelrod, 1986)—but we suggest that these "richer" definitions blur into moral judgment. In addition to many moral norms, there are thousands of nonmoral norms governing conventions of social life (Bicchieri, 2005; Machery, 2012), such as what to wear to school (Turiel, 1983). Therefore, we must ask what separates conventional norms from moral norms (Machery, 2012; Smetana, 1981).

Negative affect. One possibility is that negative affect differentiates conventional norms from moral norms (Nichols, 2002). Norm violations that garner moral condemnation are uniformly negative: giving a stranger a million dollars might violate an expectancy, but it isn't immoral. The importance of negativity is highlighted by one theory that suggests that moral judgments are "norms + feelings" (Nichols, 2002), in which strong negative affect differentiates the unconventional from the immoral. This account is consistent with research highlighting the important role of emotions in moral cognition (Haidt, 2003; Valdesolo & DeSteno, 2006). Although developmental psychology once emphasized reason over emotion in moral judgment (Kohlberg, 1969; Rest, 1975), studies find that moral judgments are often intuitive and laden with negative affect (Haidt, 2001), and that amplifying negative affect amplifies moral judgment (Horberg, Oveis, & Keltner, 2011; Inbar, Pizarro, & Bloom, 2012; Valdesolo & DeSteno, 2006).

Despite being integral to moral judgment, negative affect cannot *alone* distinguish the immoral from the unconventional as people feel negative affect and express anger when observing nonmoral norm violations (Brauer & Chekroun, 2005; Santee & Jackson, 1977). Imagine an adult spitting in their soup at a restaurant before eating it (Nichols, 2002). This act is certainly both counter-normative and disgusting, but lacks the authority-independent, objective-seeming punch of other immoral deeds.

Admittedly, people use the word "wrong" to describe soup-spitting, but this word—although popular in moral psychology questionnaires (Haidt & Hersh, 2001; Haidt et al., 1993; though see Cushman, 2008, for discussion of question wording; and Clifford, Iyengar, Cabeza, & Sinnott-Armstrong, 2015, for specifying "morally wrong")—is vague. It is "wrong" to spell morality with two "m"s, and yet few would judge this act worthy of blame or punishment.

What then separates "wrong" negative norm violation from "morally wrong" negative norm violations (Gert, 2015; Machery, 2012)? We suggest that the answer is harm.

Harm

Harm's importance to morality was suggested long ago by philosophers (Bentham, 1970/1780; Hume, 1777; Mill, 1861; Singer, 1981), and explored empirically by developmental psychologists (Piaget, 1932; Turiel et al., 1987), who revealed that both children and adults reliably use the presence of suffering victims to separate the unconventional from the immoral (Huebner, Lee, & Hauser, 2010; Turiel et al., 1987). Even 6-month old infants (Hamlin, 2012; Van de Vondervoort & Hamlin, 2016)—and some nonhuman primates (Sheskin & Santos, 2012)—use harm as a basis for social evaluation, reflecting its evolutionary importance (Haidt & Joseph, 2007). For example, rhesus monkeys avoid pulling on a chain to get food if doing so also shocks another monkey (Masserman, Wechkin, & Terris, 1964), reflecting a concern for well-being that is even more prominent in great apes (Lakshminarayanan & Santos, 2008; Warneken, Hare, Melis, Hanus, & Tomasello, 2007; Warneken & Tomasello, 2006).

Preventing harm to oneself and one's kin—that is, one's genes—was likely a key motivating factor for the evolution of morality (Haidt & Joseph, 2007; Hauser, 2006). There is nothing that impedes your genes propagating more than you or an offspring being murdered, having your livelihood stolen, having someone impregnate your spouse, getting a sexually transmitted disease, or being enslaved by a neighboring civilization. Without prohibitions against harm—whether direct or indirect—not only would genetic propagation be impaired, but cultures would likely collapse into chaos. Given its evolutionary importance, it follows that harm should be the most developmentally basic and universal psychological cause of moral judgment.

Harmful acts such as murder, rape, assault, abuse, and theft evoke powerful moral outrage (Hutcherson & Gross, 2011) and are uniformly condemned across cultures (Mead, 1961; Mikhail, 2007, 2009; Shweder, 2012). Concerns about harm also form the basis of laws (Fletcher, 1998; Holmes, 1909) which can be understood as the codification of a society's moral code (Marmor & Sarch, 2015). When new laws are passed, whether about censorship (E. N. Brown, 2014; Schein, Goranson, & Gray, 2015), gay marriage (Kennedy, 2015), or bathroom use (*Carcaño v. McCrory*, 2016), lawmakers uniformly ground them in a perceived source of harm they are intended to eliminate.

In addition, harm dominates everyday morality: a "big data" study using momentary assessments of daily life found that harm was the most frequent concern in moral judgment in both liberals and conservatives (Hofmann et al., 2014). Our own recent work also reveals that concerns about harm are the most cognitively salient—when people are asked to

generate an example of something that is "morally wrong," more than 90% offer an example of harm, such as murder or abuse (Schein & Gray, 2015).

The TDM affirms both the ultimate and psychological importance of harm in moral judgment, and harm's ability to reliably separate violations of morality from those of convention. Dyadic morality also acknowledges psychological links between harm and both norm violations and negative affect. In being immoral, dyadically harmful acts are also norm violations (Mikhail, 2007; Sripada & Stich, 2005), and generators of negative affect (Rozin, Lowery, Imada, & Haidt, 1999). Thus, the presence of harm alone can often provide all three components of immorality and fuel the flame of moral condemnation.

The importance of harm in moral judgment suggests that its nature is obvious, but unfortunately it is not. Dyadic morality suggests that harm has been largely misunderstood, and that this misunderstanding has led to broader confusion in psychology. Past operationalizations of harm have depicted harm as an objective property of an act causing physical or emotional suffering to a person, or animal—as evidenced by blood, broken bones, and tears (Graham et al., 2011; though see Turiel et al., 1987, for an argument about how harm is perceived). In this view, only acts such as murder, assault, and physical and emotional abuse qualify for the label "harm" (Graham et al., 2009). As objective properties are amenable to reasoning, harm was also seen as grounded in reason. For example, research on moral dilemmas ties considerations of harm (i.e., utilitarianism) to the operation of rational thought (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Suter & Hertwig, 2011; though see Gawronski & Beer, 2016).

In addition to assumptions of objectivity and reason, a subtler assumption in some past research is that harm is binary. While some research has certainly differentiated degrees of harmfulness (e.g., personal vs. impersonal; Cushman, Young, & Hauser, 2006), other study designs imply that harm is either present or not, such that an act is harmful or harmless (Haidt, Bjorklund, & Murphy, 2000). We suggest that all of these assumptions have led researchers to incorrectly reject harm as an explanation for diverse moral judgments.

A review of the literature suggest that harm should be redefined: rather than objective, reasoned, and binary, TDM defines harm as a *synthetic* combination of agents causing damage to patients, and suggests that harm is *perceived, intuitive*, and *a continuum*.

What Is Harm? Synthetic, Perceived, Intuitive, a Continuum

Harm Is Synthetic

The Merriam-Webster dictionary ("Definition of Harm," 2016) defines harm as "physical or mental damage or injury:

something that causes someone or something to be hurt, broken, made less valuable or successful, etc." This definition highlights two essential elements of harm within morality—causation and damage to vulnerable victims—but it misses an important third component: an intentional agent. Acts such as car accidents, unexpected falls, and natural disasters are all harmful, but they are not immoral in the same way as child abuse (Cushman, 2008). That accidental injuries and murder are both termed "harmful" is a source of confusion, similar to that surrounding the word "wrong." When TDM refers to harm, it is a very specific kind of harm: *an intentional agent causing damage to a vulnerable patient.*

This definition is *synthetic*, consisting of three elements: an intentional agent, a vulnerable patient, and a causal link of damage between them. Together, these three elements combine to yield the moral dyad of two interacting minds—a perceived agent who acts and a perceived patient who receives the act. We outline this formal dyadic definition of harm here to prevent confusion throughout the article and in discussions more broadly; when dyadic morality uses the term "harm," this is what is meant.

The dyadic structure of harm has been noted historically by philosophers (Aristotle, BC350) and legal scholars (Hart & Honoré, 1985), and reflects the dyadic structure of both action and language (R. Brown & Fish, 1983; Mikhail, 2007; Strickland, Fisher, & Knobe, 2012). Contemporary theories in moral psychology also endorse the dyadic nature of harm, whether in the structure of "ME HURT YOU" (Greene, Nystrom, Engell, Darley, & Cohen, 2004), or more complex computations that integrate causality and intentionality (Cushman & Young, 2011; Mikhail, 2007).

One popular battery of moral scenarios includes five scenarios designed to tap concerns about harm, and all five include the presence of an intentional agent causing damage to a vulnerable patient (e.g., an adult sticks a pin into the hand of a child; Graham & Haidt, 2012). Another battery includes 27 scenarios of harm and all of them follow a dyadic structure (e.g., You see a woman throwing her cat across the room for scratching the furniture; Clifford et al., 2015). These surveys—and many more scenarios used to study morality (Killen, Lynn Mulvey, Richardson, Jampol, & Woodward, 2011; Smetana, 1985)— reflect the dyadic structure of harm. Moreover, the consistent moral condemnation of these harmful violations reveals that this dyadic structure causes robust moral judgments for both liberals and conservatives (Haidt, 2007; Schein & Gray, 2015). We therefore suggest that claims that harm is dyadic—and that dyadic harm is tied to moral judgment—are uncontroversial.

Two Moral Minds

Central to TDM's understanding of harm is the kinds of minds possessed by its doer (the agent) and its recipient (the patient; Wegner & Gray, 2016). In general, people perceive mental qualities along two broad dimensions, distinguishing between

agency (doing, thinking) and experience (feeling, sensing; Gray, Jenkins, Heberlein, & Wegner, 2011). While adult humans are seen to possess both agency and experience, puppies and children are seen to possess mostly experience, and God and Google are seen to possess mostly agency (Gray et al., 2011). More colloquially, those with agency—adults, God (Gray & Wegner, 2010a), and corporations (Rai & Diermeier, 2015)—can be understood as "thinking doers," and those with experience—adults, children (Schein et al., 2015), and animals (Bentham, 1970/1780; Singer, 1975)—can be understood as "vulnerable feelers" (Wegner & Gray, 2016).

Agency is linked to moral responsibility (Eshleman, 2014; Feigenson & Park, 2006), consistent with legal doctrine that ties responsibility (*mens rea*) to the mental capacity for planning, intention, and forethought (Hart & Honoré, 1985; Shaver, 1985). Agency—being a thinking doer—is thus what qualifies one as a *moral agent* (who possesses *moral agency*). This link between mind and moral status is apparent in debates about adolescent sentencing (Steinberg & Scott, 2003) and the capital punishment of those with mental disabilities (Ellis, 2003), both of which revolve around questions of capacity for planning and intention.

Conversely, experience is tied to moral rights, consistent with philosophy that ties rights to vulnerability and the ability to suffer (Bentham, 1970/1780; Singer, 1975). Experience—being a vulnerable feeler—is thus what qualifies one as a *moral patient* (who possesses *moral patiency*), as evident from debates about abortion that revolve around a fetus's ability to feel. Pro-life advocates see a vulnerable, young child in need of protection (H.R. 356 - 109th Congress, 2005), whereas pro-choice advocates see a mindless mass of replicating cells (Benatar & Benatar, 2001). The same questions about suffering and vulnerability also exist within contemporary debates of animals rights (Bratanova, Loughnan, & Bastian, 2011; Foer, 2010), and historical debates about slavery (*Dred Scott v. Sandford*, 1856). Perceptions of patiency (or lack thereof) directly affect moral treatment (Haslam, 2006), such as when the denial of experience to Black adults predicts the endorsement of police brutality (Goff, Eberhardt, Williams, & Jackson, 2008).

It is important to note that the mere perception of suffering and vulnerability is not enough to give rise to a robust moral judgment. Instead, one must *care* about the vulnerable mind via empathy (Baron-Cohen, 2011; Eisenberg & Miller, 1987). Psychopaths can identify that others can feel—though less so than normal people (Gray et al., 2011)—but lack empathy for their suffering (Blair, 1995). Empathy is not a "cold" cognitive state of simply recognizing suffering; even from a young age, others' suffering evokes emotion (Eisenberg, Spinrad, & Sadovsky, 2006), and motivates altruistic behavior (de Waal, 2008). Newborn infants recognize the crying of others (G. B. Martin & Clark, 1982; Simner, 1971), and pre-linguistic children show inclinations toward helping others in distress (Warneken & Tomasello, 2006). Importantly, children show more empathy for suffering caused by legitimate moral wrongs, than for "cry-babies," suggesting that recognizing interpersonal harm is a unique impetus for morality (Leslie, Mallon, & DiCorcia, 2006).

Predicting Moral Condemnation With Minds and Causation

A dyadic understanding of harm allows us to predict which acts should be judged as most harmful and therefore as most immoral. Moral judgment is proportional to the agency of agents, the experience of patients, and the clarity of causation between them; acts with obviously intentional agents who cause obvious damage to obviously vulnerable patients should seem both most harmful and immoral. Consistent with this idea, people will robustly condemn the CEO of a corporation (obvious thinking doer) for taking a 2-year-old (obvious vulnerable feeler) and hitting her across the face (causing obvious damage). On the contrary, people will not generally condemn a 1-year-old who squishes an ant, because toddlers lack agency and ants lack experience.

In addition to the general mental capacities of agents and patients, moral judgments also hinge on the agent's specific intentionality and the patient's specific suffering within a given act. Someone capable of deep reflection may kill another person, but they will be blamed less if they were of unsound mind at the time, or acted accidentally (Malle et al., 2014; Roberts, Golding, & Fincham, 1987). Likewise, an agent who targets someone generally sensitive to pain but fails to harm them (attempted murder) will be punished less than one who actually kills them (Young & Saxe, 2010).

Harm is not merely the presence of an intentional agent and suffering patient but also involves a causal act linking them together (Cushman, 2015; Cushman et al., 2006). The clearer this causation, the more obvious the harm, which is why moral judgments are attenuated when causality is unclear (Fincham & Roberts, 1985; Fincham & Shultz, 1981). The distinction between action and inaction (or omission and commission) arises from differences in perceived causality (Baron & Ritov, 2004). Poisoning an opponent by recommending a salad containing peanuts is judged as more immoral than not stopping the opponent from ordering it himself (Spranca, Minsk, & Baron, 1991) because failing to speak is causally opaque (DeScioli, Bruening, & Kurzban, 2011).

A clear real-world example of the importance of clear causation comes from hiking drug prices. If Company A wants to raise prices of a life-saving medication, they are condemned if they do it directly. On the contrary, if they do it through selling it to an intermediary, it seems less immoral, because the causation (and intention) are indirect (Paharia, Kassam, Greene, & Bazerman, 2009). The ambiguity of causation also explains why mafia kingpins put layers of people (and ignorance) between themselves and dirty deeds.

The power of dyadic elements is made clear by the competing strategies of lawyers during court proceedings. Prosecutors typically argue for the intentionality

and premeditation of the defendant, the vulnerability and suffering of the plaintiff, and the clear damage caused. In contrast, defense attorneys argue that no damage was caused, that the plaintiff is not vulnerable and suffered little, and that their client did not intend any wrongdoing (sometimes by virtue of insanity; Roberts et al., 1987). The success of these arguments upon juries arises from the fact that harm is subjective, with the elements of agency, experience, and causation each a matter of perception.

Harm Is a Matter of Perception

In Jonathan Foer's (2010) book "Animals," he interviews two turkey farmers. One of them—a factory farmer—sees farming as a matter of pure economics, and fails to see either harm or immorality when birds are kept confined in tiny cages with their beaks and claws seared off. The other—a man who maintains his own personal farm—sees farming as a moral issue, and abhors the idea of harming his birds, in whom he sees a rich mental life. That two adult men can view the same issue and come to such different conclusions suggests that harm is in the eye of the beholder. Dyadic morality allows for the subjectivity of harm because its synthetic elements—agency, vulnerability, and causality—are all perceived.

Questions of mind are essential for judgments of harm and immorality, but the nature of other minds is ambiguous (Hyslop, 2015). You can be certain of your own mental states (Descartes, 1641), but those of others can never be directly experienced, and so are ultimately inaccessible (Avramides, 2001). This "problem of other minds" means that we can never be certain whether someone possesses the capacity for agency (i.e., intention and planning) when they act. For example, when two 12-year-old girls tried to murder their friend during a sleepover (The Associated Press, 2016), they claimed that—at the time—they lacked an appreciation of right versus wrong. Of course, there may be an objective fact about the lucidity of their minds during the attempted murder, but these facts only affect moral judgment via others' perceptions.

As ambiguous as agency may be, questions of patiency/experience are even more opaque because feelings of pain and suffering are internal sensations (A. Smith, 1812). Even with tears and screams, whether someone (or something) is legitimately experiencing pain is up for debate. For example, a 20-week-old fetus will twitch and grimace when poked in the heel (Benatar & Benatar, 2001), but does this reaction demonstrate real pain? The same concerns appear with adults, such as when people question whether a victim of sexual assault "really" suffered or is just trying to get attention or money. These perceptions of suffering are shaped by various factors such as the victim's race (Hoffman, Trawalter, Axt, & Oliver, 2016), the perceived deservingness of the pain (Weiner, 1993), and even the number of people in pain (Jenni & Loewenstein, 1997).

Just as with mind, questions of causality are ambiguous. As Hume remarked long ago, we can never directly observe causation but only infer its presence from the co-occurrence of two events (Hume, 1740). This inference of causation can be clear when it involves physical events such as a cue hitting an eight ball (Michotte, 1963), but causation in human events is more opaque (Wegner & Wheatley, 1999), especially when events are complex (Phillips & Shaw, 2015). For example, when a town's water supply is poisoned 30 years after the closing of a mine, is it because of natural occurring arsenic or because of corporate malfeasance? Both may cause suffering, but only the latter is "harmful" in a dyadic—and moral—sense.

Perceived harm also plays a role in acts that lead only probabilistically to suffering, such as drunk driving or negligence (Malle et al., 2014; J. W. Martin & Cushman, 2016a). Typically, when someone drives after a few drinks, they arrive home safely and do no damage. However, sometimes drunk drivers will crash their vehicle, destroying people's property and sometimes killing the innocent. People respond strongly to these cases of clear suffering but also to acts of potential harm because they can entertain counterfactuals—thinking "what if . . . ?" (Spellman & Gilbert, 2014). Some have critiqued dyadic morality as being incapable of accounting for the moral condemnation of negligence or recklessness (Pizarro, Tannenbaum, & Uhlmann, 2012), but this criticism only holds if people do not link these acts to potential harm—and they do.

Others have critiqued dyadic morality for being unable to account for the victim blaming (Niemi & Young, 2016) frequently observed after sexual assault (Niemi & Young, 2014; Suarez & Gadalla, 2010). However, dyadic morality explains this phenomenon as ambiguity about the "true" identity of the agent and patient (Hester & Gray, 2017). While the physical perpetrator and victim within rape is clear, the ultimate (i.e., mental) agent is a matter of perception, with some perceiving the sexual assaulter as ultimately the victim of the sexual assaultee's capriciousness or mixed signals. Cases of victim blaming—and sympathizing with the perpetrator—are instances of "dyadic reversal" in which perceptions of mental agency and patiency are exactly opposite the physical structure of the act (Hester & Gray, 2017). Consistent with classic accounts of attribution (Weiner, 1993), the more a victim seems to be a responsible agent, the less she is seen as suffering, and the more a perpetrator seems to be a suffering patient (e.g., after receiving harsh punishment), the less he is seen as responsible (a hydraulic relation called "moral typecasting"; Gray & Wegner, 2009).

The perceived nature of harm is reinforced by anthropological accounts revealing that although norms concerning harm are universal, the specifics of who can do harm (agents), who can receive it (patients), and how it can be done (causation) are culturally constrained by different understandings of mind and causality (Mead, 1961; Shweder, 2012). Cultural learning informs perceptions of

who can intend, what constitutes causality, and what can be harmed. Holding livestock morally responsible seems preposterous now, but in medieval Europe, animals such as pigs were seen fit to stand trial (Oldridge, 2004). Cultures can also constrain empathy, such as via propaganda campaigns of dehumanization that define others as enemies rather than people (Haslam, 2006; Staub, 1992). Even perceptions of causation are embedded in cultural worldviews. If witchcraft exists, then suddenly the mere chanting of incoherent syllables can cause moral harm. Even just within the United States, Protestants are more likely than Jews to believe that beliefs in one's heart can be the source of harmful outcomes (Cohen & Rozin, 2001).

In the terms of Turiel and colleagues (1987), there are different "informational assumptions" (p. 189) about the nature of the world that shape perceptions of harm. Ascribing causal power to witchcraft allows for evil to be perpetrated through thoughts or glances (e.g., the evil eye), opening these behaviors up to condemnation (Cohen & Rozin, 2001; Pronin, Wegner, McCarthy, & Rodriguez, 2006). Cultures who emphasize an immortal human soul also believe that the dead can be harmed—and so moralize funeral rites (Rachels, 1986; Shweder, 2012). For example, Brahman Indians believe it is highly immoral for an eldest son to eat chicken after his father's funeral because he is responsible for purifying his father's "death pollution" through eating a vegetarian diet. By eating chicken, he undermines this process and condemns his father's soul to eternal suffering (Shweder, 2012). Similarly, the Hua of New Guinea believe that the blood of individual members of a community are intertwined, and so defiling one's own body threatens communal well-being (Meigs, 1984; cited in Haidt, Rozin, Mccauley, & Imada, 1997).

These perceptions of causation and vulnerability are often motivated (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996), and can be amplified through anthropomorphism (seeing mind in nonhumans; Epley, Waytz, & Cacioppo, 2007) and diminished through dehumanization (stripping mind from humans; Haslam, 2006). In the case of the turkey farmers described above, it would be distressing for the factory farmer to believe that he was complicit in turkey genocide, and so he strips away mind from the birds and sees his techniques (e.g., pumping birds full of antibiotics) as merely mechanistic. Indeed, empirical research has found that after eating animals, people are more likely to deny the mind to similar animals, and the mere categorization of an animal as food undermines mental trait attribution (Bastian, Loughnan, Haslam, & Radke, 2012; Bratanova et al., 2011). Similarly, interviews of prison personnel found that executioners exhibited the greatest tendency for dehumanization of prisoners relative to staff supporting the death-row inmates, and noninvolved prison guards (Osofsky, Bandura, & Zimbardo, 2005).

The subjectivity of harm is perhaps revealed best in modern political discourse, such as when one person sees gay marriage as an expression of love, and another sees the destruction of families, souls, and the minds of children (Adam, 2003; Anderson, 2013; Comer, 2012). Compared with liberals, conservatives are more likely to be religious and nationalistic (Layman, 2001), explaining in part why they might see both souls and America as vulnerable to harm (Haidt, 2012). Political disagreements about harm also appear in moral debates about transgender bathroom use (*Carcaño v. McCrory*, 2016), masturbation (Laqueur, 2004), and pornography, which the Republican National Committee has labeled "a public health crisis" (Peters, 2016). One might argue that one side of each debate is wrong, and that there is an objective fact about harm, such that pornography either causes harm or not. However, we suggest that claims of "objective" harm are misplaced, at least as they concern moral psychology.

The field of moral psychology has long accepted the perceived nature of morality (Schein, Hester, & Gray, 2016), but for some reason, has implied that harm is objective—such as when researchers reassure participants that acts of consensual incest are objectively harmless (Haidt et al., 2000). Such assumptions contradict not only the experience of political debates but also the anthropological accounts upon which modern moral psychology is built (Shweder, 2012; Shweder, Mahapatra, & Miller, 1987). We suggest that there is no firm basis to accept morality as perceived but not harm—especially when they always seem to go together. Of course, one could argue that moral judgments and perceptions of harm are different in kind, with moral judgments being intuitive and perceptions of harm being reasoned rationalization. We evaluate this claim next.

Harm Is Intuitive

Considerations of harm loom large in moral reasoning and moral rhetoric whether it concerns pornography, prostitution, genetically modified foods (GMOs), stem cell research, gun rights, gay rights, or free speech (Haslam, 2016; Kennedy, 2015; Lukianoff & Haidt, 2015; Schein & Gray, 2016). Despite (or perhaps because of) the clear evidence of harm's prominence in moral reasoning, some scholars conclude that harm has a limited role in intuitive moral judgment (Haidt, 2012). However, this division between intuition and reason is a false dichotomy as many psychological elements used in reasoning are also highly intuitive (Kruglanski & Gigerenzer, 2011).

Research from diverse areas suggest that perceptions of harm are intuitive. Children too young to engage in conscious rationalization—and too young to use language—use harm in their social evaluations (Hamlin & Wynn, 2011). Research using high-density event-related potentials (ERPs; Decety & Cacioppo, 2012) and implanted brain electrodes (Hesse et al., 2015) reveal that adults can not only see harm within milliseconds but also differentiate its varieties (Decety & Cacioppo, 2012). Our own research reveals that judgments of harm can happen rapidly and without conscious deliberation (Schein & Gray, 2015). In fact, when people are asked to

judge issues (e.g., pornography) on both immorality and harmfulness, they respond just as fast—if not faster—to questions about harm (Schein & Gray, 2015). This is exactly the result one would expect if considerations of harm were not only intuitive but also fed into moral judgments.

The intuitiveness of harm may be controversial, but so once was the intuitiveness of morality (the debate is actually ongoing; Smetana & Killen, 2008). Early developmental accounts of morality emphasized the importance of reason in moral cognition (Kohlberg, 1969; Rest, 1975). However, a revolution in social psychology revealed that—in nonmoral domains—reasoning often plays second fiddle to intuition and affect (Nisbett & Wilson, 1977; Zajonc, 1980). Inspired by this revolution, Haidt (2001) and Greene and colleagues (2001) revealed that moral judgments are often rapid and affectively negative reactions that arise with relatively little thought (Haidt, 2001). Of course, reasoning can play an important role in shaping moral judgment (Paxton & Greene, 2010; Smetana & Killen, 2008) and conversation (Mercier, 2011), even if its presence is relatively infrequent.

The main evidence that harm is a product of reasoning and not of intuition comes from moral dumbfounding research (explored later) and moral dilemmas (i.e., trolley problems), which are constructed to pit utilitarian reasoning (kill one to save five) against deontological intuitions (don't kill; Greene et al., 2001). As considerations of harm have been equated with utilitarianism, they have been casted as opposing the intuitions of deontology. This logic is faulty, however, as deontological intuitions are potentially more about harm than are utilitarian reasoning—after all, deontological intuitions are reactions against *killing*.

Furthermore, moral dilemmas frequently confound utilitarianism and deontology with action and inaction (Gawronski & Beer, 2016). For example, the utilitarian option in trolley problems almost always involves committing an action (e.g., push a man), whereas the deontological option involves inaction (e.g., don't push a man). Research also reveals that deontological and utilitarian impulses do not conflict but are independent, a fact revealed by the technique of process dissociation (Conway & Gawronski, 2013).

More broadly, trolley problems are engaging precisely because they pit reason against intuitions, which is atypical in moral judgment. People do not agonize over whether murder is wrong because of intuitions or reasons because it is both—it feels wrong to murder, and there are also good reasons not to do it. It is worth stating explicitly that perceiving harm seldom involves rational, utilitarian cost-benefit analysis. People are lousy moral utility calculators (Sheskin & Baumard, 2016), and generally not consequentialist thinkers (Baron, 1994). Instead, harm is an intuitive perception—one that is affective and often motivated, and which may or may not cohere to an external objective reality of the immediate or absolute number of people killed.

One common rational for denying the intuitiveness of harm comes from a strict—and heavily contested (Helion &

Pizarro, 2015; Kruglanski & Gigerenzer, 2011; Van Bavel et al., 2015)—dual-process model of the mind in which the processes of intuition are qualitatively distinct from those of reasoning (Greene, 2013; Haidt, 2012). While this dual-process account aligns with historical conceptions of the passions versus reasons (Descartes, 1641; Hume, 1740), it does not represent modern understanding of decision making (Lindquist & Barrett, 2012). Instead, intuitions and reasoning often operate in tandem, in which initial intuitive judgments are modified by conscious reasoning (given sufficient capacity or motivation; Haidt, 2001). We suggest that the same is true with harm—initial intuitive perceptions of harm can be modified by additional conscious reasoning.

We acknowledge that in courts of law, people do deliberately process questions of intent, suffering, and causality to arrive at judgments of blameworthiness (Hart & Honoré, 1985; Malle et al., 2014). However, each of the three dyadic elements—and therefore harm—are also processed intuitively in both adults and children. Even if considerations of intention can involve sophisticated reasoning, adults intuitively perceive the presence of intentional agents (H. C. Barrett, Todd, Miller, & Blythe, 2005) and infer intentions (Rosset, 2008)—and infants effortlessly see goals behind even minimal actions (Woodward, 2005). Even if reactions to others' suffering can be dampened though motivated emotion regulation (Cameron & Payne, 2011), both adults and children intuitively perceive the suffering of others (Decety & Jackson, 2004; Eisenberg & Miller, 1987; Preston & de Waal, 2001), which automatically generates empathic concern (Eisenberg et al., 2006). Even if causation can be considered explicitly (Jones & Nisbett, 1972), adults and children also intuitively perceive causation in physical (Leslie & Keeble, 1987; Michotte, 1963) and social events (Hartshorne, 2014).

Overall, both harm itself and the elements composing harm are perceived intuitively, even if reason can strengthen or weaken these perceptions. The same is true about moral judgments, which arise intuitively and yet may be shaped by reason. Indeed, the predictions of dyadic morality concern intuitions of both harm and morality: *intuitive* perceptions of harm should predict *intuitive* judgments of immorality, along a continuum.

Harm Is a Continuum

Is watching professionally produced pornographic films harmful? People feel very strongly about the answer to this question. One could argue that women in the adult film industry are being exploited (Dworkin, 1981), but one could also argue that women use sex to empower themselves and make a reasonable living (Miller-Young, 2013). However, even staunch opponents of pornography would admit that the typical pornographic film with consenting adults is less harmful and less immoral than child pornography. In adult pornography, there is ambiguity about whether the film stars
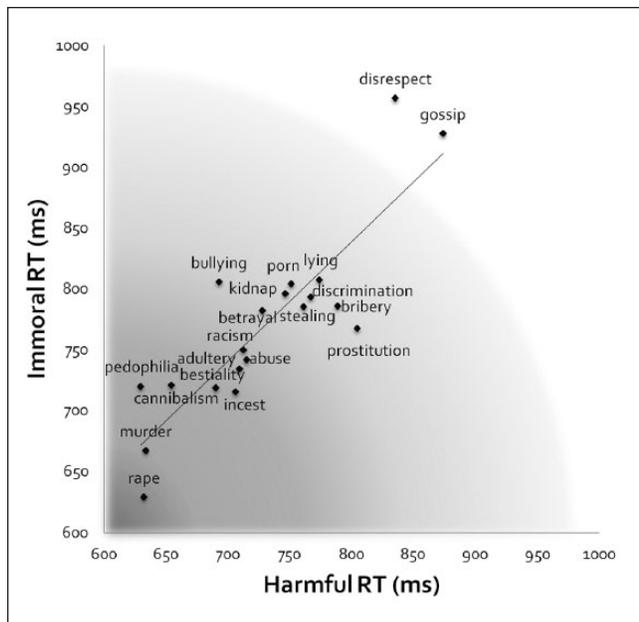
Figure 3. Response times for categorizing an act as immoral and harmful. They reveal an intuitive continuum of harm that corresponds to an intuitive continuum of immorality (Study 4; Schein & Gray, 2015).

are agents or patients, but no ambiguity exists with child pornography as children are clear victims of exploitation.

Dyadic morality suggests that moral judgments are aligned on a continuum (or gradient) of harm (see Figure 2). At the most severe end are acts in which intentional agents, vulnerable patients, and the causation of damage are very easy to perceive, such as prototypically immoral—and evolutionarily threatening—acts such as murder, rape, and abuse. At the least severe end are acts in which it is extremely difficult (but not technically impossible) to perceive any dyadic elements, such as meditating alone outside on a summer's day. In between these ends of black and white is a vast expanse of gray, in which the salience of agents, patients, and causation—that is, harm—is ambiguous and therefore subject to disagreement.

One need only examine legal sentencing to see that morality is a continuum, in which people are executed for premeditated murder, imprisoned for defrauding investors, and merely fined for speeding. This legal continuum is explicitly tied to the degree of harm each of these acts seems to involve, with aggravated assault causing more obvious suffering than fraud. The graded nature of harm also exists in people's minds. Our own work reveals that when people make rapid judgments of the harmfulness of various immoral acts—ranging from murder to gossip—their reaction times form a clear gradient (Schein & Gray, 2015; Study 4). Murder and rape are rated as harmful at ~600 ms, followed by adultery and abuse at ~700 ms, followed by prostitution and lying at ~800 ms, followed finally by disrespect and gossip at ~850

ms. Importantly, this intuitive gradient matches the intuitive gradient of immorality, and exists even when controlling for general negativity (see Figure 3).

This continuum of harm is facilitated by the continua of an agents' perceived general agency, and a patient's perceived general experience (Gray et al., 2011; see also S. T. Fiske, Cuddy, & Glick, 2007; P. Robbins & Jack, 2006; Sytsma & Machery, 2009), as well as their specific agency and experience within an act, and the causality connecting them (Cushman, 2015; J. W. Martin & Cushman, 2016b; Murray & Lombrozo, 2016; Phillips & Shaw, 2015; Spellman & Gilbert, 2014). When infants die of heatstroke after being forgotten, parents are blamed, but less so than when infants are beaten to death. When bartenders serve a drunk patron who then drives home and kills someone, they are blamed, but less so than a bartender who directly poisons someone (Cushman, 2015; J. W. Martin & Cushman, 2016b; Murray & Lombrozo, 2016; Phillips & Shaw, 2015; Spellman & Gilbert, 2014). When a would-be murderer misses his victim who nonetheless dies while racing from the scene, he is blamed, but less so than the directly successful shooter (Cushman, 2008; Robinson & Darley, 1995). As none of these examples are controversial, we suggest that the gradient of harm—and its synthetic elements—is uncontroversial.

What Is the Moral Dyad? A Fuzzy Cognitive Template

A continuum of harm means that the dyadic definition of (A → P) does not describe a strict philosophical definition of immorality with a clear boundary, but instead a fuzzy psychological template (Figure 1). This fuzziness is not a dodge, but instead a fact of life: It is impossible to draw a firm in/out line around any rich human concept, as Wittgenstein (1953) discovered long ago when trying to determine the authoritative definition of "games."

Consistent with Wittgenstein's experience, modern cognitive science reveals that concepts are not instantiated by a list of binary inclusionary and exclusionary criterion, but instead by a set of graded synthetic criteria whose combined presence determines a continuum of concept categorization (Medin, 1989; E. E. Smith & Medin, 1981). For example, the stereotype of "African American" includes a number of features including skin tone, facial features, speaking style, manner of dress, occupation, and socioeconomic status (Eberhardt, Goff, Purdie, & Davies, 2004; Gaertner & McLaughlin, 1983). A target is judged stereotypically African American to the extent that these features are collectively salient, such that hip-hop artist 50 Cent seems more African American than President Obama, who in turn seems more African American than Bill Gates.

The moral dyad is a similarly fuzzy template—and is therefore consistent with volumes of work on social cognition (S. T. Fiske & Taylor, 1991). Its synthetic elements are

an intentional agent, vulnerable patient, and causation of damage, and acts that better match these combined criteria—by being more saliently dyadic—are more robustly categorized as harmful and thus as immoral. More succinctly, an act is immoral to the extent of its perceived "dyadicness." A graphical representation of the fuzzy dyad is therefore not a circle, but instead a circular gradient (see Figures 1-3).

In the (dark) middle of the template are acts that are maximally dyadic such as murder, rape, child abuse, and animal slaughter—the kind of acts used in various "harm" scenarios (Graham & Haidt, 2012). Moral theories agree that both harm and immorality exist in this dark, central, obviously dyadic area (Haidt et al., 2015). Likewise, theories agree that neither harm nor immorality likely exist in the light obviously nondyadic areas. However, some theories assume that harm is sharply circumscribed and ends with direct physical harm, while TDM advocates for gradations of harm, which can include intuitive perceptions of *some* harm in acts of spiritual destruction and societal dissolution.

Understanding morality as a continuum offers a deflationary perspective on research arguing for cognitive differences across moral scenarios (Young & Saxe, 2011). One set of studies revealed that considerations of intention were less important for acts involving sex than those involving violence (Young & Saxe, 2011). Specifically, the difference in moral condemnation between accidentally versus intentionally committing incest was smaller than the difference in moral condemnation between accidentally versus intentionally committing murder. This result was interpreted as revealing distinct cognition between the moral judgment of sex versus violence (i.e., purity and harm), but it failed to consider the importance of relative severity. People rate murder as *more* harmful and immoral than incest (Gray & Keeney, 2015b). As intention is an element of harm, we would expect intention to matter more for the more harmful murder—and it does. Said another way, all these results reveal is that the elements of immorality are relevant in moral judgment to the extent that something is immoral. More broadly, when making claims about different *kinds* of immorality, one must first control for different *amounts* of immorality.

An intuitive continuum of perceived harm suggests that the question researchers should ask themselves—and participants—when conducting moral psychology is not "is it immoral?" or "is it harmful?" but "*how much* does *x* seem immoral?" and "*how much* does *x* seem harmful?"

The Key Prediction of Dyadic Morality

Given an updated understanding of harm as a synthetic, perceived, intuitive continuum, we now revisit the main prediction of dyadic morality, that acts are immoral to the extent that they are harmful. TDM suggests that moral judgments are made according to the intuitive perception of synthetic harm—intentional agents causing damage to vulnerable patients (A → P). This suggests that very harmful acts should be judged as very immoral, that somewhat harmful acts should be judged as somewhat immoral, and that minimally harmful acts should be judged as minimally immoral.

The alignment of harm and immorality is reasonably easy to test. One may simply obtain intuitive perceptions of harm and see if they predict intuitive judgments of immorality. Importantly, this harm is not merely badness or sadness, but perceptions of damage caused to a vulnerable entity at the behest of an intentional agent. Consistent with these predictions, both our research and those of others reveals that perceptions of harm are robustly tied to moral condemnation (Gray, Schein, & Ward, 2014; Royzman et al., 2015; Schein & Gray, 2015; Schein, Ritter, & Gray, 2016)—research that we explore throughout this article.

One point of clarification is that while some perception of harm is typically necessary for an act to enter the (fuzzy) moral sphere, harm is not the *only* relevant consideration. Once an act is moralized, then cultural norms or emotions (e.g., disgust) can amplify moral convictions and moral outrage (Wisneski et al., 2009; Wisneski & Skitka, 2017), which may be tied to certain behaviors (Brandt, Wisneski, & Skitka, 2015). In addition, we acknowledge that we are describing a statistical relation and not a philosophical absolute, and so there may be exceptions to the predictive power of harm.

Nevertheless, dyadic morality predicts that any exceptions to dyadic morality are rare, unstable, and maintained only with effortful reasoning. In fact, we suggest that most apparent exceptions to dyadic morality are not exceptions at all; rather, they are cases in which harm is misunderstood to be an objective, reasoned fact rather than a perceived intuitive continuum (Haidt & Hersh, 2001; Scott et al., 2016). Nowhere is harm more misunderstood than in moral dumbfounding and the condemnation of "harmless" wrongs (Haidt et al., 2000).

Are There Harmless Wrongs? No: Speaking for Moral Dumbfounding

Dyadic morality argues that harmless wrongs—although logically possible—should be psychologically rare. Consistent with this idea, unless there is some mitigating feature (e.g., self-defense),[2] people never seem to say that something is both obviously dyadically harmful (A→ P) and obviously permissible, as revealed by judgments of scenarios involving direct canonical harm (Turiel, 1983). However, one program of research suggests that people will sometimes endorse the wrongness of ostensibly harmless acts—moral dumbfounding (Haidt et al., 2000).

In one classic study, participants were presented with an "objectively" harmless act such as secret, loving, consensual incest and asked to judge its morality (Haidt et al., 2000). Participants uniformly judged this act as immoral, despite experimenter's explicit reassurances of their harmlessness. When participants were later asked to explain their moral

judgments, they were rendered speechless—a phenomenon labeled "moral dumbfounding."

Moral dumbfounding is among the most discussed effects in contemporary moral psychology; although it is unpublished, this study has garnered more than 200 citations as of 2016. Most interpret these findings as revealing the irrelevance of objective harm in moral judgments (Russell & Giner-Sorolla, 2013), and the existence of a harm-independent purity mechanism (Haidt, 2001, 2007; for arguments in this tradition, see Scott et al., 2016). However, we suggest that moral dumbfounding actually provides powerful evidence in *support* of the harm's role in moral judgment, when harm is properly understood.

Perceived Harm in "Harmless" Wrongs

The most obvious limitation with moral dumbfounding is that it conflates objective harm with perceived harm. Researchers may have "carefully written" the stories "to be harmless" (Haidt et al., 2000, p. 8), but this does not mean that participants share their perspective. In one published set of studies, adults in both Philadelphia and rural Brazil were asked about the morality of consensual incest (Haidt et al.,1993). The data revealed that 45% of adults judged this act as universally immoral, and 64% judged this act as deserving of punishment. Importantly, other measures also revealed that people perceived this act as harmful, with 59% of participants identifying victims damaged from this act. The close alignment between perceptions of harm and judgments of immorality are exactly what dyadic morality predicts.

In a more recent set of studies, Royzman and colleagues (2015) recreated the traditional dumbfounding paradigm. But this time, they measured whether participants actually believed the acts to be harmless, for example asking, "how believable do you find that Julie and Mark's (who are siblings) having sex with each other will have no bad consequences . . . " (Royzman et al., 2015, p. 300). In this design, participants simply rejected the set-up of the incest vignette, holding on to the belief that incest has negative real-world implications. The incredulity of participants persisted in individual interviews, even after researchers argued in favor of incest's harmlessness. It is not hard to explain this persistence of perceived harm. After all, incest is reliably harmful in almost every single context except in moral psychology scenarios.

We find it perplexing that scholars inspired by anthropology have failed to acknowledge these perceptions of harm. Imagine if emotion researchers claimed to have designed scenarios which were "objectively" angering or saddening and yet failed to elicit these emotions in participants. It is unlikely that such claims of "objective emotionality" would be accepted. We believe that same evidentiary standards should be used in moral psychology; if participants deny that these scenarios are harmless, then we should accept their experience as legitimate. Given the anthropological roots of modern moral psychology, the recommendation that we take seriously participant perceptions should be uncontroversial—even if (or especially when) they conflict with the designs of researchers.

Intuitive Harm in Dumbfounding

Also perplexing is how scholars will accept a priori that perceptions of morality are intuitive but not those of harm. When people are presented with consensual incest, their immediate—that is, intuitive—response is to call it harmful (Royzman et al., 2015). Intuitive responses are hard to dispel with conscious reasoning, and harm is no exception; even with experimenters forcefully arguing with them, participants cannot shake their perceptions of harm. In fact, participants are only "rendered dumb" when they are banned from reporting what they see as the basis of their moral judgment—harm. We suggest that preventing participants from mentioning harm misleads moral psychology from identifying the true underlying causes of moral judgment.

As an analogy, consider someone who fears flying because it seems dangerous to them. You could argue that flying is actually quite safe—and they might rationally agree with you—but their perceptions of danger are intuitive and therefore resistant to reason (Mayer, Merckelbach, & Muris, 2000). Flying may be objectively harmless, but subjective perceptions of potential harm *do* legitimately cause people's fear (Seligman, 1971). Moreover, if you asked people with flight phobias to tell you why they feared flying while banning them from mentioning feelings of danger, they would be rendered dumb—but this tells us nothing deep about the causes of their fear.

Our own studies reveal additional evidence for the intuitive persistence of harm in "harmless wrongs." In one study, we asked participants to rate the immorality and harmfulness (i.e., the presence of victims) of three kinds of acts: those that were unambiguously harmful (e.g., murder), unambiguously harmless (e.g., taking the bus), and *ostensibly* harmless acts such as rubbing feces on a Bible (Gray, Schein, & Ward, 2014). Half of the participants were given ample time and told to think carefully (to tap "reason"), and half were given a tight time restrain and were told to respond with their gut (to tap "intuition").

The study revealed that perceptions of harm were *higher* when people were forced to respond rapidly. These initial, intuitive perceptions of harm are not merely metaphorical or symbolic (though an appeal to symbolic harms can certainly occur later; Gutierrez & Giner-Sorolla, 2007), as people see physical pain and children's misery in wrongdoing (Gray, Schein, & Ward, 2014). Consistent with dyadic morality, harm is tied to immorality when assessed intuitively, with reason diminishing this link.

A Continuum of Harm in Dumbfounding

Classic moral dumbfounding studies claim that objectively harmless acts are seen as immoral—a statement that

collapses continua into dichotomies. People certainly do condemn acts like consensual incest, but we seldom ask "compared to what?" Having loving sex with your sibling may be more immoral than fudging your taxes, but it is likely not as bad as perpetrating genocide. When researchers claim that consensual incest (or eating your dead dog, or having sex with a chicken, etc.) is immoral, they are implicitly making homogeneous the entire moral sphere, as if there is only one level of morality. The same homogenizing is found in claims of "harmlessness" (Haidt, 2001). Even if we ignore the robust intuitive perception of harm in consensual incest, we can still agree that, "objectively speaking," having sex with your sibling involves more harm (or potential harm) than watching a movie with them, but less harm than stabbing out their eyes. When immorality and harm are seen as continua, then the claims of moral dumbfounding are much less impressive: acts which are somewhat harmful are somewhat immoral.

Not only do acts vary on their moral severity, they also vary on how typical they are in everyday life. Morality evolved to guide our everyday interpersonal interactions and facilitate cooperation (Haidt & Kesebir, 2010), and yet moral psychology has thousands of citations to papers with bizarre moral acts such as frozen-chicken-sex, masturbating with kittens, eating dead dogs, selling your soul, getting a plastic surgery tail, and doing performance art that involves rolling around in urine (Eibach et al., 2009; Haidt, 2001; Haidt & Hersh, 2001; Helzer & Pizarro, 2011). We suggest that these weird and ambiguous acts do not best reveal the nature of moral cognition (Gray, 2014; Gray & Keeney, 2015a, 2015b).

Building a model of moral judgment based upon bizarre acts such as consensual incest and chicken-sex is like building a model of mammals based upon the platypus. These bizarre scenarios are certainly interesting, but—like platypuses—are so precisely because they are unrepresentative of their broader category. The vast majority of immoral acts in daily life are uncontroversially harmful such as causing someone physical or emotional injury, denying someone their rights, treating someone unfairly, and betraying someone's confidence (Hofmann et al., 2014). It is these more typical, harmful acts which should be the focus of the field, at least if we want our theories to apply to the real world. We should resist a "platypus moral psychology."

What About Other Values? They Are Transformations and Intermediaries of Harm

Although direct physical or emotional harm is most powerfully and universally immoral, people moralize values including purity, loyalty, honesty, industriousness, benevolence, and rationality (Haidt, 2007; Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011; Schwartz & Bilsky, 1987). The importance of these values varies from culture to culture, as

does their moral status (Rai & Fiske, 2011). For example, conservatives appear to moralize obedience to authority more than liberals (Frimer, Gaucher, & Schaefer, 2014; Stenner, 2005), and liberals moralize the purity of the environment more so than conservatives (Frimer, Tell, & Haidt, 2015).

Past research has misinterpreted dyadic morality as opposing pluralism (Graham et al., 2013), but nothing could be further from the truth: Dyadic morality embraces diverse norms and cultural variability in moral judgment. However, TDM suggests that harm—perceived, intuitive, and dyadic—is integral to moralization, and that values are moralized to the extent that they are associated with harm. Central to this idea is a conceptual distinction between values and morality; people hold numerous values, but not all are moralized.

Consider punctuality. Many individuals in industrialized countries believe it is important to arrive on time, a view not uniformly shared by those in developing countries. However, even industrialized countries reveal wide variation based upon culture and context: if an American tells you to arrive at 8:00 a.m. for an interview, you shouldn't be a minute late, but if a Spaniard tells you that a party starts at 8:00 p.m., you might arrive closer to 10:00 p.m. Despite the pluralism of this value, punctuality is not necessarily a moral value. Showing up late for an interview is foolish, but not worthy of moral condemnation. In cases where punctuality is moralized—showing up late to your daughter's wedding or failing to mail your child's college application on time—it is because of the harm caused by such tardiness. Admittedly, such harm is small compared with that of murder, but such relative differences are consistent with a continuum of harm, and with TDM, which suggests that values will be moralized to the extent that they involve perceived harm.

Values and Norms

Earlier, we suggested that morality involves three components: norms, negative affect, and harm. Values share psychological similarities to norms, an idea supported by their parallel definitions. Schwartz (1999) defined values as "conceptions of the desirable that guide the way social actors . . . select actions, evaluate people and events, and explain their actions and evaluations" (p. 24), and Cialdini et al. (1991) defined injunctive norms as "rules or beliefs as to what constitutes morally approved and disapproved conduct" (p. 1015). Equating values to norms allows us to understand when they should be moralized—when their violation generates negative affect, and is perceived as harmful.

It is uncontroversial that the violation of norms or values induces negative feelings (Brauer & Chaurand, 2010; Brauer & Chekroun, 2005), but the role of harm in moralization of diverse values is more contentious. We suggest this controversy stems from past research that conflates values with moralization. MFT suggests there are five specific values
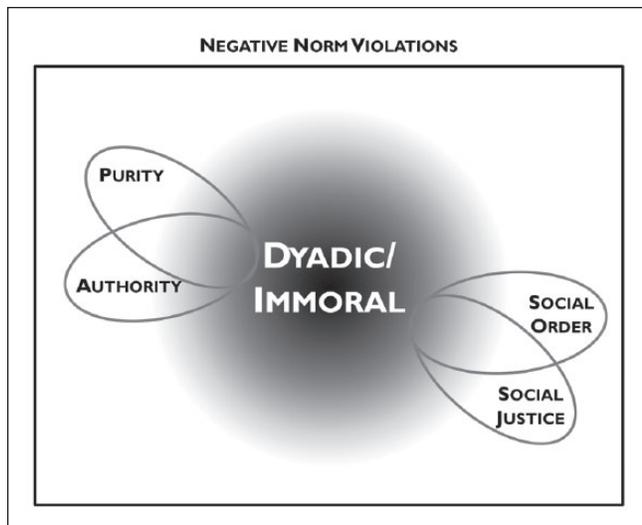
Figure 4. Dyadic morality suggests that negative norm violations are moralized to the extent that they involve perceived harm. As important values can be understood as norms, dyadic morality allows moral pluralism through the variability of norms across culture. Values are moralized to the extent that they are intuitively bound to harm.

(harm, fairness, loyalty, authority, and purity) which are intrinsically moralized—although only some people have these values "activated" through cultural learning. This idea of intrinsically moralized values has difficulty explaining how people can understand—and even hold—these values without moralizing them. For example, even liberal professors recognize the importance for children (and their graduate students) to respect their authority, but may not moralize it to the same degree as conservatives.

It is also problematic to restrict the set of moralized norms to a specific number, as it fails to capture how people moralize hundreds of other values such as security, self-direction, achievement, power, or hedonism (Schwartz & Bilsky, 1987). Recent research has even found that rationality is a moral value for some individuals (Stahl, Zaal, & Skitka, 2016). One could argue that the diversity of moralized norms boils down to this subset of five or six, but one could then argue that these five boil down further—perhaps to concerns about harm.

Harm Moralizes Norms

Once we separate values from morality, we can ask what distinguishes moral values from nonmoral values (see Figure 4). This is analogous to the question of what distinguishes moral norms from nonmoral norms, and we provide the same answer as we did earlier: harm. Violations of values seem immoral when they seem harmful—when harm is understood as dyadic, perceived, intuitive, and continuous. Linking the moralization of values to harm provides a domain-general

ingredient that can be applied to any value, not just a narrow subset.

Of course, some values may be more or less easily tied to harm—often because of evolutionary considerations—and are therefore more or less easily moralized. The value of protecting the lives of your children is clearly tied to harm (A → P), and so it is robustly moralized. In contrast, the value of chastity is less clearly tied to harm, and so moral disagreement surrounds this value. Importantly, disagreement does not surround ratings of chastity *per se*, as most people appreciate that teenage sex violates the value of chastity. Rather, the disagreement is about whether the value of chastity is a moral value, which we suggest is tantamount to the question of whether violations of chastity seem harmful. Consistent with this idea, moral opponents of teenaged sex argue that violating chastity leads to diseases, destroys teens' ability for emotional intimacy, and damages their morality (Fullinwider, 1994).

In line with TDM, research reveals that perceptions of harm predict moral judgments concerning other values. In one study, moral condemnation of purity scenarios (e.g., rolling around in urine, selling your soul; Graham & Haidt, 2012) was predicted by perceptions of harm ($r = .70$; Gray & Keeney, 2015b). In another study, moral condemnation of GMOs—hypothesized to be purity-related (Scott et al., 2016)—was best predicted by perceived harm (Gray & Schein, 2016). Indeed, perceptions of harm accounted for 30 times the variance than did feelings of disgust—despite past predictions regarding the unique link between disgust and purity (Horberg, Oveis, Keltner, & Cohen, 2009). We acknowledge that these ratings are self-report and correlational, but this study nonetheless suggests that the people who view GMOs as harmful are also the ones who view it as a moral value.

The predictive power of harm regarding GMOs is even more striking when you focus upon so-called "moral absolutists." These are participants who state that they would be opposed to GMOs no matter what the benefits, and so—the authors hypothesized—should moralize purity independently from any link to harm (Scott et al., 2016). However, these "moral absolutists" were even *more* likely to tie their moral condemnation of GMOs to perceptions of harm versus disgust (Gray & Schein, 2016).

Perhaps the strongest evidence for the intuitive link between harm and the moralization of values comes from a study that used an implicit association test. In this study (Schein & Gray, 2015; Study 7a), participants categorized words as either Immoral or Not Immoral. The Not Immoral words were *forget, procrastinate*, and *boring*, and the Immoral words all tapped violations of the value of "purity": *incest, bestiality*, and *prostitution*. Participants also categorized words as either Harmful (*victim, harmful*, and *dangerous*) or Disloyal (*disloyal, unfaithful*, and *unpatriotic*). Past research argues that the purity is part of a "binding moral

foundation" which includes loyalty. This suggests that the more impure acts are seen as immoral, the more closely they should be linked to loyalty. Conversely, dyadic morality predicts that the more impurity seems immoral, the more impurity should be linked to harm. Consistent with dyadic morality, the more that people moralized the value of purity, the more they tied its violation to harm over loyalty—with no differences across politics. In other words, across both liberals and conservatives, it was the association with harm that best predicted whether purity was seen as a moral concern.

Taxonomies and Intermediaries of Harm

Harm can predict the moralization of diverse values because the dyadic template allows for different *transformations* of harm, depending on contexts and cultures. Like origami, moralized values involve transformations of harm—not through folding, but through flexibility in who is identified as agents and patients, and what is identified as a damaging act. For example, impure acts can involve damage to the body, the future self, the soul, or society, and the more salient this perceived damage, the more robust these acts are moralized.

One question is what happens if people moralize values in themselves? For example, some people decry freedom as a good in itself (Iyer, Koleva, Graham, Ditto, & Haidt, 2012). We predict that the moralization of these abstract values are still tied to harm, and that these values are intermediaries of harm. By intermediary, we are referring to a concept (e.g., purity) that is seen as a vulnerable entity in its own right. For example, in explaining the immorality of a purity violation such as a widow eating "hot" food, one of Shweder's Indian participants noted that the act not only impacts her deceased husbands spirit (a direct harm), but the act will lead her to "lose her sanctity" (i.e., harm her purity), a consequence that in turn leads to her tangible suffering (Shweder et al., 1987, p. 44). Dyadic morality therefore allows for two links to harm: the direct perception that an act is harmful, and the indirect perception that an act destroys a value—which then causes direct harm. This prediction is currently untested, but is easy to evaluate: for example, do conservatives intuitively perceive harm when "loyalty" is violated?

Dyadic morality embraces moral diversity across cultures and the importance of moral values, but is so far silent upon which exact values are embraced by a culture. There are other theories, however, that catalog these values, including MFT (Haidt, 2007), Relationship Regulation Model (Rai & Fiske, 2011), and the Model of Moral Motives (MMM; Janoff-Bulman & Carnes, 2013). These taxonomies serve an important role because they provide intuitive categories for moral diversity and facilitate moral discussions (Graham et al., 2013; Rai & Fiske, 2011). Most consistent with dyadic morality is the MMM (Janoff-Bulman & Carnes, 2013), which divides moral content into six cells depending on the

identity of the moral patient (future self; other; group) and whether harm is being perpetrated or prevented.

Regardless of one's favored taxonomy, we caution against confusing taxonomic groups with cognitive natural kinds (Lindquist, Siegel, Quigley, & Barrett, 2013). Describing the language of values in anthropological accounts, on Twitter or in Congress is valuable, but cannot reveal the structure of moral cognition any more than text analyses of Rolling Stone can reveal the structure of musical cognition. Moreover, "naming is not explaining" (Gawronski & Bodenhausen, 2015, p. 18): giving a set of moral violations an intuitive label does not automatically reveal the deep roots of moral cognition. For example, when MFT ascribes political disagreement about sex education to differences in the importance of "purity," it does not explain the cause of this debate. Instead, it merely repackages it into a tautology because purity is *defined* as moral concerns about sex and religion. In other words, MFT suggests that debates about the morality of sex are due to differences about the morality of sex. Because these tautological labels furnish the illusion of explanation, they disguise the complex combination of factors underpinning moral conflict, such as religiosity, historical accident, and competing perceptions of harm (Schein & Gray, 2016).

Dyadic morality thus does not seek to name moral disagreement, but instead to explain it. It suggests a simple answer to the question of which values a society moralizes: those associated with harm.

Changing Harm, Changing Moralized Values

Dyadic morality also helps explain why moralized values differ between cultures and over time (Schein & Gray 2016; 2017). Changing perceptions of harm parallel changes in these values, as the value of sexual purity clearly illustrates. In the ancient world, an act such as masturbation was not only permitted, but sometimes even celebrated because it was thought to increase fertility and strength (Laqueur, 2004). Masturbation only became the subject of moral outrage during the Enlightenment because of changing perceptions of harm. In the early 1700s, an anonymous author published the pamphlet "Onania; or, The Heinous Sin of Self Pollution, and all its Frightful Consequences . . . " (Laqueur, 2004). The pamphlet had widespread international distribution, and ignited moral panic.[3] Masturbation lost its moral tinge with the normalization of sex (Kinsey, Pomeroy, Martin, & Sloan, 1948) and the awareness of the countervailing harm of sexual shaming (Day, 2013).

Cultural variations of perceived harm may sometimes seem mysterious but are sensible once we consider a people's evolutionary and cultural environment (Asch, 1952; Douglas, 1966; Rachels, 1986). For example, cultures see sexual promiscuity as harmful when the prevalence of sexually transmitted infection is high (Tybur, Lieberman, Kurzban, & DeScioli, 2013), or see disloyalty as harmful when cultures face annihilation in warfare (Gelfand et al.,
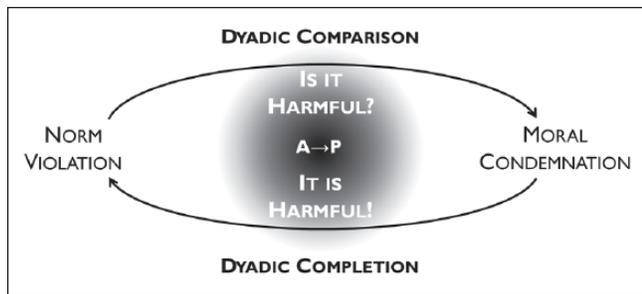
Figure 5. The complementary causal processes of dyadic comparison and dyadic completion. Norm violations that are perceived as harmful are morally condemned (dyadic comparison), and once condemned, these acts are perceived as more harmful (dyadic completion).

2011). Of course, perceptions of harm—like beauty or any other cultural product—are also affected by stochastic factors that result in random drift and arbitrary differences between cultures. As with evolutionary drift, random variation in perceived harm (and morality) may then be exapted to serve a useful function, such as binding cultures together around shared but arbitrary judgments (A. P. Fiske, 2000; Greene, 2013).

In sum, dyadic morality suggests that harm separates the moral from the nonmoral, whether it concerns acts or values. The role of harm in mapping out moral content is provided by the mechanism of moral judgment.

Mechanism

Questions of *what* are inextricably bound to questions of *how*. In the content section, we have suggested that dyadic harm is central to morality. In this section on mechanism, we explore the processes by which perceptions of harm shape—and are shaped by—moral judgment. We review the complementary dynamic processes of dyadic comparison and dyadic completion, and the dyadic loop that their combination yields (see Figure 5). We then discuss broader debates about the mind—modularity versus constructionism—and how they impact moral psychology. We conclude by examining moral emotions and disgust.

Although these discussions focus upon mechanisms of harm, we again acknowledge the importance of norms and negative affect in constructing morality. Harm-based processes of moral judgment are likely initialized only once a norm violation is noticed, and the strength of the final moral judgment hinges on its associated negative affect, which is typically integral or related to an act (Wisneski & Skitka, 2017) but can also be generated through incidental manipulations (Schnall, Haidt, Clore, & Jordan, 2008; though see Landy & Goodwin, 2015; Johnson et al., 2016).

Importantly, as dyadic acts (e.g., murder) are often intrinsically negative norm violations, these factors are somewhat—but not fully—contained within the moral dyad.

Questions of how exactly norms and affect are psychologically instantiated are beyond the scope of this article, but are reviewed in moral psychology (Haidt, 2003; Sripada & Stich, 2005) and in nonmoral psychology (L. F. Barrett & Russell, 2014; Bicchieri, 2005). Indeed, we suggest that nonmoral psychology is an excellent guide for how we make moral judgments.

Making Nonmoral Judgments

Moral judgments are unique in their severity, conviction, and implications (Skitka et al., 2005; Skitka et al., 2015), and so it is tempting to treat moral mechanisms as equally special. Despite the appeal of searching for the "morality" region in the brain that automatically categorizes moral violations (Derbyshire, 2010), a decade of research suggests that no single, specific region exists (Greene & Haidt, 2002; Young & Dungan, 2012). Rather, research reveals that moral judgments involve the same domain-general cognitive processes and brain regions used in other decision-making paradigms (Cushman & Young, 2011; Van Bavel et al., 2015; Miller & Cushman, in press). We therefore suggest that moral judgments are made similar to nonmoral judgments, and that asking people "Is *x* act immoral," is analogous to asking "Is *y* animal a mammal?" or "Is *z* person African American?" All three involve *categorization*, asking whether a specific example is representative of a category or concept.

Categorization has been heavily studied in cognitive psychology (Medin, Wattenmaker, & Hampson, 1987; Murphy, 2004), and although researchers disagree on some details, there is widespread consensus on how these decisions are made—and not made. Categorizations are *not* made with strict verbal definitions containing a list of necessary and sufficient conditions (Wittgenstein, 1953). Although this idea was historically popular, it fell out of fashion as philosophers and psychologists realized its inability to capture the gradient (or continuum) of categorization—how some examples can be better (e.g., horses) or worse (e.g., platypuses) examples of a category (e.g., mammals; Medin, 1989). Categorization decisions are also *not* made with modules—encapsulated and innate mechanisms (Sperber, 1994)—as modules are too rigid to account for contextual effects and variations in categorization (Fodor, 1985). While the strict assumptions of modularity can be relaxed, we will later argue that weak modularity lacks falsifiability, and only furnishes the illusion of explanation.

Rather than involving definitions and modules, research reveals that categorization decisions are made by comparing examples to cognitive templates (Rosch, 1978). These templates are synthetic constructions that combine the most important, salient, and typical features of category members. For instance, in "birds," this template contains the element of feathers, flight, and beaks; in "African American," this template contains elements of skin tone, facial features, dress, speech, and socioeconomic status. Templates can be thought

of as either a single prototype (Rosch, 1978), or a set of exemplars (E. E. Smith & Medin, 2002), but the important point is that they synthesize features.

Categorization decisions are then made by comparing examples—rapidly and intuitively—to this synthetic template, with better matches resulting in more robust categorization decisions. This explains why a robin is a "better" bird (i.e., more bird like) than an ostrich, and a golden retriever is a "better" dog than a Puli. Literally hundreds of studies in cognition and social cognition support this mechanism of categorization in nonmoral judgment (Rosch, 1978; E. E. Smith & Medin, 1981; Taylor, 2003). We suggest that such "template comparison" also operates within morality.

How Do We Make Moral Judgments? Through Dyadic Comparison and Direct Learning

Dyadic morality suggests that moral judgments are typically made via template comparison with a dyadic cognitive template (A→P). That is, moral judgment occurs via *dyadic comparison*, with more dyadic acts being judged as "more immoral." Dyadic comparison can be captured by the phrase *the more something seems harmful, the more it seems immoral*.

The process of template comparison in nonmoral judgment is revealed by several phenomena: a continuum of categorization judgments, a continuum of categorization speed, and a continuum of categorization accessibility—each of which is predicted by the template features. Studies reveal that the same is true in moral judgment, with a dyadic template predicting the phenomena of moral judgment.

Evidence for Dyadic Template Comparison

A continuum of judgments. Categorization decisions exist along a continuum between obvious members and obvious nonmembers, with the level of categorization predicted by how much an example matches the template (Rosch, 1978). In morality, the more dyadic an act, the more it should be judged as immoral. In other words, more harmful acts should be (cognitively speaking) "better" examples of the concept of immorality. Consistent with this idea, experimental tests that manipulate the presence of harm find that perceptions of harm predict the severity of moral judgments (Schein & Gray, 2015).

In a different study, we asked participants to rate the harmfulness, disgustingness, weirdness, and immorality of a variety of harmful and disgusting acts, and found that across actions, the perception of harm was the best predictor of immorality ratings, even when controlling for disgust and weirdness (Schein, Ritter, & Gray, 2016). Moreover, increasing the salience of each dyadic element—intention, causation, and suffering—all increase the strength of perceived harm and moral condemnation (Ames & Fiske, 2013;

Cushman, 2008; Cushman et al., 2006; Jenni & Loewenstein, 1997; Malle, 2006; Malle et al., 2014; J. W. Martin & Cushman, 2016b; Young, Cushman, Hauser, & Saxe, 2007). Although many of these studies rely on tasks that allow for deliberative thinking, they also rely on between-subjects designs, where participants are presumably unaware that the researchers are manipulating elements of the vignettes (e.g., Ames & Fiske, 2013).

A continuum of speed. With nonmoral judgment, examples that better match the category template are not only more robustly categorized as category members but are also more quickly categorized (Rosch, 1978). The same continuum of categorization speed exists in morality. In one study, we gave participants a speeded reaction time task, in which they were asked to categorize 60 different acts on their immorality, harmfulness, and unpleasantness (Schein et al., 2015). Consistent with predictions, we found that, for both liberals and conservatives, the more harmful the act, the more quickly it was categorized as immoral, even when controlling for general negativity (see Figure 3; Schein et al., 2015).

A continuum of accessibility. Cognitive science also reveals that category examples which best match the template are most accessible. If you tell someone to think of a mammal, they mention "dog" more often than "platypus." We tested the accessibility of moral acts by asking 100 Mturk participants to "list an act that is morally wrong. Write down whatever comes to mind first." Results reveal that—for both liberals and conservatives—more than 90% of the recalled acts were dyadic, such as murder, stealing, adultery, or abuse (Schein & Gray, 2015).

A Causal Connection?

There is clearly a connection between an act's perceived harmfulness (A→P) and the moral condemnation it engenders. However, one question is whether this association is correlational or causal. Perhaps acts are judged as immoral based on some other criteria, and judgments of immorality subsequently evoke matching perceptions of harm. As we will see, moral judgments do cause perceptions of harm—via dyadic completion—but harm is also causally connected to moral judgment, as recent experiments in our lab reveal.

In this set of studies (Hester et al., 2017), participants were presented with nonsense actions taken from linguistics research, such as *John gished Mary*. Given that these actions are not intrinsically immoral (or even meaningful), the only source of moral condemnation is their surrounding context. Across studies, we manipulated the presence of a dyadic context (*John gished Mary* vs. *John gished*), the presence of intention (*John intentionally gished Mary* vs. *John accidentally gished Mary*), and the presence of suffering (*John gished Mary, who cried* vs.

John gished Mary, who laughed). Consistent with TDM, the addition of a dyadic context, intention, and suffering, all caused increased moral condemnation—all without a specific moral act. Importantly, sentences were read and judged under time pressure, suggesting it was *intuitive* perception of harm causing *intuitive* moral judgments.

One challenge to the causal role of harm comes from the idea that acts can be directly learned as immoral without invoking harm. Studies with children reveal that while they do learn morality through experience with harm (Blair, 1995; Nucci, Turiel, & Encarnacion-Gawrych, 1983), adults also seem to teach children that some acts are "just wrong." Direct learning from parental testimony occurs in morality (Rottman et al., 2015)—and in many other categories (Gelman, 2009; Harris, 2015)—but children do not blindly accept all testimony equally (Koenig, Clément, & Harris, 2004). With morality, past research suggests that children are most likely to moralize actions which innately evoke empathy or outrage (Blair, 1995)—that is, those involving harm (Rottman et al., 2015).

Parents also often infuse moral testimony with harm, such as when religious parents condemn masturbation by linking it to angels dying or Jesus crying. This co-occurrence is effective because, as initial studies reveal, moralization—the transformation of an issue from nonmoral to immoral (Rozin, 1999)—is most robust via harm (Rottman et al., 2015). Strikingly, even ostensibly harmless norms can be moralized through harm-based testimony, as revealed by a series of studies in which 7-year-old children were presented with novel actions (e.g., someone "covering their heads with sticks"). When children were told that arbitrary actions were harmful, they judged them as immoral, even 3 months later (Rottman et al., 2015).

The causal power of harm in moral acquisition allows for moral learning, and also suggests which moralized norms are most likely to persist intuitively—those most robustly tied to harm. While children (and adults) can be told that any act is immoral, TDM suggests that acts remain intuitively immoral to the extent that they are intuitively associated with harm, even if such intuitions of harm defy "objective" reasoning. The causal role of intuitive harm is also supported by research on psychopaths, who can both learn norms and feel negative affect, but whose lack of intuitive empathy to others' suffering prevents them acquiring intuitive morality (Blair, 1995).

In sum, harm both amplifies and directly causes moral judgment, even in cases where acts seem "objectively" harmless. We acknowledge that perceptions of harm need not be consciously considered when making moral judgment, especially with frequently encountered acts; most people just "know" that murder and rape are immoral, just as they just "know" that dogs are mammals. But such rote associations do not disprove the existence of a causal cognitive template in either case, especially when the speed and strength of these associations remain exquisitely predicted by harm.

How Do Moral Judgments Affect Perceptions? Through Dyadic Completion

Cognitive templates shape "bottom-up" judgments of how much examples represent a concept; they also shape "top-down" perceptions of those examples after those judgments have been made (Read, Vanman, & Miller, 1997). If dyadic comparison can be summarized as "*the more something seems harmful, the more it seems immoral*," then dyadic completion suggests the complementary process of "*the more something seems immoral, the more it seems harmful*."

Although the most severe immoral acts are obviously dyadic, less obviously dyadic acts may nevertheless be categorized as somewhat immoral—whether because of rote learning or because they are dyadic "enough" to trigger a weak moral judgment. As in other categorization decisions, making a moral judgment initiates an unconscious coherence process that aligns perceptions of a specific example with the broader conceptual template (Clark et al., 2015; Thagard, 2002)—which leads ambiguously dyadic acts to seem even more dyadic. This process is called dyadic completion because a somewhat incomplete dyad (lacking a clear agent, patient, or causation) is "completed" by perceiving the missing agent, patient, or causation (Gray & Schein, 2012).

The presence of evil agents and suffering patients (A   P) compels *causal dyadic completion*, the perception of a causal link between them. For example, a drug dealer rushing home to hide cocaine is ascribed more causal responsibility for a car crash than someone rushing home for a more innocuous reason (Alicke, 1992). It is also the reason why evil CEOs seem more causally responsible than good CEOs for harming the environment (Knobe, 2003), and why evil thoughts seem to cause the suffering of others more than good thoughts (Pronin et al., 2006).

Similarly, isolated suffering patients (→P) compel *agentic dyadic completion*, the perception of intentional agents to account for their suffering. This is why people blame—and sue (Sunstein, Kahneman, & Schkade, 1998)—individuals (Knobe, 2003), corporations (Stout, 2013), God (Gall, 2004; Gray & Wegner, 2010a), and even animals (Oldridge, 2004) after accidents and natural disasters. Our studies reveal that the experience of suffering on both a small and large national scale prompts belief in an intentional God (Gray & Wegner, 2010a).

Perhaps the most important form of dyadic completion is *patientic dyadic completion* in which norm violations completed by intentional agents (A→) compel people to see moral patients harmed by those acts (DeScioli, Gilbert, & Kurzban, 2012; Gray, Young et al., 2012; Gray, Schein, & Ward, 2014). Research finds that people even see suffering in events which one could argue are "objectively" harmless, such as burning an American flag, eating a dead dog, grave desecration, and homosexuality (DeScioli et al., 2012).

Given the vulnerability of children, dyadic completion often activates perceptions of the suffering of children. Indeed, "Think of the Children" is a catchphrase for moral campaigns against pornography (Pierce, 2001), for and against gay marriage (Kennedy, 2015), transgender bathroom use (Philipps, 2016), and masturbation (Day, 2013). Although liberals may see this harm as mere rhetoric, both anthropological accounts (Shweder, 2012) and experimental studies reveal these acts are legitimately perceived to have concrete victims (Clark et al., 2015; DeScioli et al., 2012; Gray, Schein et al., 2014).

Importantly, our own research suggests that perceptions of harm occur automatically (Gray, Schein et al., 2014). In one study, we primed participants with ostensibly harmless acts such as desecrating a bible, strange masturbation, and necrophilia, and then presented participants with images of children, who were perceived as expressing more suffering but not more boredom (a negative nonharm control). In a related study, participants who had just read about burning an American flag or having sex with a store-bought chicken (Haidt et al., 1993) saw unrelated injuries as more painful, but did not see Chinese pictographs as more unpleasant (a negative control). Such automatic perceptions of harm can help support moral acquisition by deepening moral judgments. Indeed, when 7-year-old children were taught that arbitrary norm violations were "just wrong," they intuitively linked those violations to harm, which helps to entrench them (Rottman et al., 2015).

Dyadic completion is likely the reason why people are seldom moral absolutists. Very few people (beyond Kant, 1780) separate immorality from perceived harm and claim that something is wrong but beneficial (Clark et al., 2015). For example, those who think GMOs are morally permissible see them as solving problems of hunger, whereas those who think GMOs are immoral see them as harming children and the environment (Gray & Schein, 2016). These perceptions of harm occur automatically and intuitively (Gray, Schein et al., 2014), but can be augmented by reasoning (DeScioli et al., 2012). In particular, we suggest that reason transforms general intuitions of harmfulness to the identification of specific victims, agents, or causal pathways (Clark et al., 2015; Ditto & Liu, 2011; Liu & Ditto, 2013). Like all perceptions of harm, this identification of dyadic elements is sensitive to context and culture (Schein & Gray, 2016).

How Are Moral Judgments Extended and Entrenched? Through the Dyadic Loop

Dyadic comparison is a bottom-up process that transforms perceived harm into judgments of immorality, and dyadic completion is a top-down process that transforms judgments of immorality into perceived harm. Together, these complementary processes form a feedback loop that can amplify moral judgments as perceptions of harm and immorality

mutually reinforce each other, a process called the *dyadic loop*. Such feedback cycles are well documented in cognition (Spivey & Dale, 2004), especially in the field of nonlinear dynamics (Vallacher, Read, & Nowak, 2002), which would consider the moral dyad an "attractor state" (Corneille, Hugenberg, & Potter, 2007). This attractor state exerts cognitive gravity that pulls ambiguous acts toward it, synchronizing and amplifying judgments of harm and immorality. This cognitive gravity can help explain the general "creep" of immorality and harm, through which people are seen to deserve greater rights and protection (Haslam, 2016; Schein & Gray, 2016)—perhaps helping to explain moral progress (Pinker, 2011).

This iterative process should not be understood as a slow process of discrete steps (Schein & Gray, 2014). The mind does not work in discrete chunks via boxes and arrows (Rumelhart & McClelland, 1986; Spivey & Dale, 2004; Thelen, 1996) but instead via parallel constraint satisfaction processes through which percepts—that is, harm and immorality—mutually reinforce each other (Read et al., 1997). This dynamic dyadic loop offers a simple explanation for the dynamics of moralization (Schein & Gray, 2016): once a person or a community sees an act as somewhat harmful—whether for functional reasons or randomly—that act will seem somewhat immoral, which will make it seem more harmful, and then more immoral, and then more harmful, then more immoral, and so on (see Figure 6). Consider smoking or vegetarianism (Rozin, 1999; Rozin, Markwith, & Stoess, 1997). What initially is a matter of preference can rapidly transform to an issue of morality once some notion of harm is introduced—especially to vulnerable moral patients (e.g., children suffering via secondhand smoke or animals suffering in factory farming).

The feedback cycle between harm and immorality can also help explain moral polarization (Haidt, 2012), in which minor disagreements between groups expand into major moral chasms—especially when group dynamics are added to the mix (Gray, Rand, et al., 2014). Immorality judgments shape factual beliefs about the world (Liu & Ditto, 2013), and so once initial perceptions of harm cause moral judgments, people then perceive the harmfulness of issues differently, such as when moral opponents of GMOs deny scientific findings regarding their benefits (supplementary material, Scott et al., 2016). Perceptions of harm are both the cause and the consequence of moral disagreement, consistent with a dynamic causal model advanced by TDM.

How Does Dyadic Morality Relate to MFT? It Argues Against Its Claim of Modularity

The emphasis of harm in dyadic morality puts it at odds with MFT (Haidt, 2007). MFT argues that people have a set of five innately prepared cognitive modules (Haidt & Joseph, 2007), one for each different kind of moral content: harm,
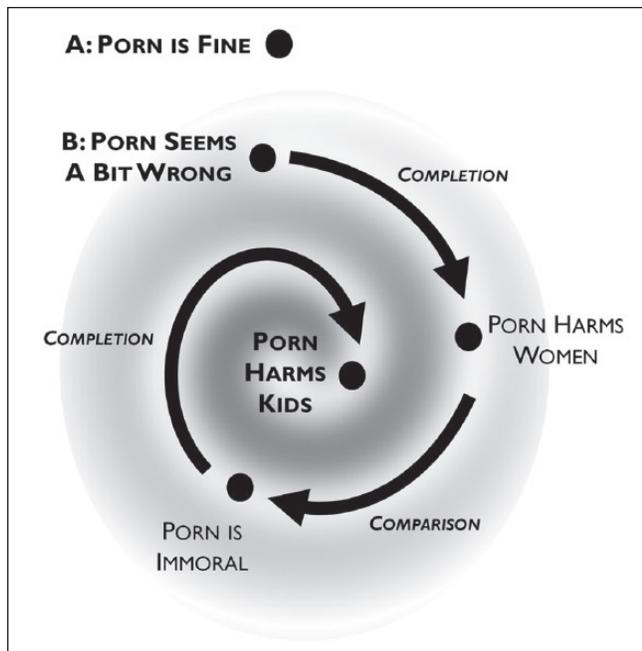
Figure 6. The dyadic template exerts cognitive gravity, making moral violations seem both more harmful and more immoral via the *dyadic loop*, characterized by the iterative and complementary maxims "what seems harmful is wrong" and "what seems wrong is harmful." This feedback loop allows for moralization and facilitates deepening moral polarization.

fairness, loyalty, authority, and purity. For example, MFT holds that having sex with a dead chicken is judged as immoral via a purity module, which makes its perceived harmfulness completely irrelevant (Haidt & Hersh, 2001).

As we reviewed above, there is ample recent evidence for the intuitive and causal role of harm within these acts. However, it is important to consider this debate more deeply, as dyadic morality and moral foundations endorse competing viewpoints about the broader nature of the human mind.

Points of Agreement

First, both TDM and MFT agree that harm can cause moral judgment—that perceptions of intention, causation, and suffering are sufficient for moral judgment. In TDM, this is accomplished through comparison with a dyadic template, and in MFT, this is accomplished through the activation of an encapsulated harm mechanism. Second, both theories agree that harm—intentionally caused suffering—is likely the most important, frequent, and universal moral consideration (Haidt et al., 2015).

Third, both theories agree upon four claims set out by MFT, that morality is *intuitive, pluralistic, culturally learned,* and *innate* (Graham et al., 2013). TDM embraces all four of these criteria: both morality and harm are intuitive, and are also compatible with pluralism via diverse moralized values,

which are best understood as transformations or intermediaries of harm. Conceptions of harm are also shaped by culture learning through different informational assumptions about mind perception, causation, and which entities are vulnerable to damage (Cohen & Rozin, 2001; Crimston et al., 2016; Haslam, 2006; Turiel et al., 1987).

Understandings of harm are also innate (Govrin, 2014). Before a baby can walk, they show signs of harm-based sociomoral judgments (Bloom, 2004; Van de Vondervoort & Hamlin, 2016), enabled through the innateness of intention (Carey, 2009; Woodward, 1998, 2005), causation (Leslie & Keeble, 1987), and suffering (G. B. Martin & Clark, 1982; Simner, 1971)—all processes that emerge early in development.

Despite these points of agreement, there remain strong areas of contention between TDM and MFT—namely the endorsement of constructionism or modularity.

Modularity

Dyadic morality and moral foundations are grounded in different views of the mind. MFT explicitly endorses modularity, in which different mental events are tied to the operation of different cognitive modules, defined by MFT as follows:

> Modules are like little switches in the brains of all animals. They are switched on by patterns that were important for survival in a particular ecological niche, and when they detect that pattern, they send out a signal that (eventually) changes the animal's behavior in a way that is (usually) adaptive. (Haidt, 2012, p. 123)

This definition reflects a "strong" claim of modularity, in which modules are cognitive mechanisms distinct from each other in terms of inputs, activation, and outputs. This feature of distinctness is essential to modular claims because they allow for *causal* claims: a specific subset of stimuli (e.g., kind of moral violations) activates one certain module (e.g., purity foundation) which causes one certain behavior or experience (e.g., purity-related moral judgment).

Reinforcing this strong modular view are analogies between moral foundations and taste buds, each of which are biologically distinct, triggered by a specific set of food chemicals, and linked to one specific psychological experience (e.g., a taste of sweetness, or saltiness; Haidt, 2012). The very act of naming moral concerns "foundations" (vs. genres) emphasizes the idea that these five moral concerns are reified as distinct mechanisms, as does the claim that modules are "little switches in the brain."

Testing Strong Modularity

Claims of strong modularity have two large benefits: explanation and falsifiability. By grounding phenomena in the activation of distinct functional mechanisms, modules allow these phenomena to be explained. When someone asks "why
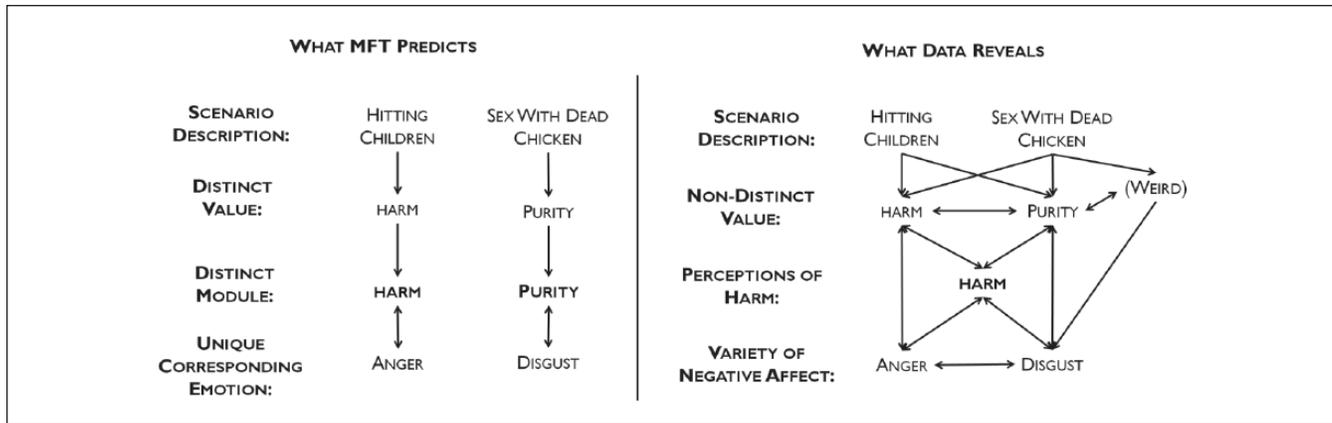
Figure 7. MFT argues for a set of isomorphic mechanisms, in which a specific moral scenario activate a distinct moral foundation—defined as a cognitive module (i.e., "little switch in the brain"; Haidt, 2012, p. 123)—which is tied to a distinct emotion. However, empirical data reveals little to no distinctness between concerns, as harm and purity are highly correlated at $r > .87$ (Gray & Keeney, 2015b), and are not tied to distinct emotions (Cameron, Lindquist, & Gray, 2015). The pervasive role of harm across morality is consistent with dyadic morality.
Note. MFT = Moral Foundations Theory.

does a peach taste sweet?" we can say "because sugar ($C_6H_{12}O_6$) binds with the 'sweet' taste bud." Distinct modular moral foundations—if present—would also allow us to explain moral judgment. When someone asks "why does someone think consensual incest is immoral?" we can say "because incest activates the 'impurity' foundation." Strong modules also allow falsifiability because it is relatively easy to test claims of distinctness and specificity.

Testing distinctness. To test distinctness, one can assess whether the psychological experience ostensibly tied to one module is independent from that of a different module. With taste buds, putting sugar on the tongue (an input) activates "sweetness," but does not trigger the experience of saltiness, and so we can call them distinct. Testing such distinctness in MFT may be difficult, as acts could activate more than one module (e.g., romantic cheating activates loyalty and harm), but thankfully, MFT researchers have designed a bank of scenarios (Graham & Haidt, 2012; Graham et al., 2009; Haidt et al., 1993) custom-designed to tap only one "foundation." These scenarios thus represent the MFT equivalent of pure salt and pure sugar; the question is whether the judgments they activate are actually distinct.

The answer is no. Across multiple studies, there is no evidence for distinctness when using these custom scenarios (Gray & Keeney, 2015b; Gray, Schein et al., 2014; Schein & Gray, 2015). In MFT discussions, purity and harm are often cast as maximally distinct (Haidt, 2001; Horberg et al., 2009), and yet their judgments are highly overlapping, with correlations greater than .87 between them (Gray & Keeney, 2015b). Other foundations are similarly highly overlapping, with the ratings of all 5 "foundations" often forming a single scale, α = .89 (Study 5; Schein & Gray, 2015). Most strikingly, although MFT researchers have designed the Moral

Foundations Questionnaire (MFQ; Graham et al., 2011) to illustrate the distinctness of moral domains, it actually reveals the opposite, with correlations of .88 between authority and loyalty, .72 between harm and fairness, and .80 between purity and authority (Figure 3; Graham et al., 2011). These correlations make claims of distinct mechanisms all but impossible, especially considering that these interfoundation correlations are higher than the intrafoundation reliabilities (i.e., loyalty is more correlated with authority than with itself). See Figure 7.

To make this lack of distinctness explicit, consider again the MFT analogy of taste buds (Haidt, 2012). If taste buds were as nondistinct as moral foundations, then even pure sugar would taste sweet, salty, bitter, sour, and umami—and so would pure salt. These empirical results suggest that while MFT may be a convenient taxonomy of overlapping values, it certainly does not capture a set of innate moral "taste buds" or cognitive mechanisms.

We acknowledge that some studies do reveal some apparent differences in moral judgments across content areas, but these findings are problematic. For example, studies arguing for a special link between purity and character (Uhlmann & Zhu, 2013; Young & Saxe, 2011) confound purity with weirdness when they use bizarre scenarios. When participant-generated examples of impurity (e.g., pornography, prostitution) are used to eliminate these confounds, any apparent differences disappear (Gray & Keeney, 2015b).

Testing moral emotions—Especially disgust. Another key claim of modularity is that specific "basic" emotions are linked to distinct moral foundations, such as purity to disgust, and harm to anger (Horberg et al., 2009). However, recent experiments and a review of the literature found no evidence for such morality-emotion specificity (Cameron, Lindquist, &
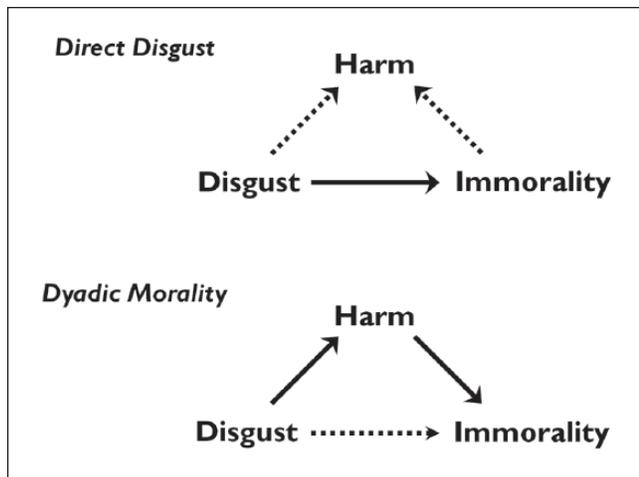
Figure 8. Two competing predictions about the mediational structure of harm, disgust, and moral judgment. "Direct disgust" argues that disgust directly causes moral judgments, with perceptions of harm only rationalizing these judgments (Haidt & Hersh, 2001). Dyadic morality argues that harm most proximately predicts immorality. Despite the long-standing popularity of the "direct disgust" hypothesis, research using a broad sample of disgusting stimuli reveals more support for dyadic morality: ratings of harm mediate the link between disgust and immorality, even in "objectively harmless" purity violations (Schein, Ritter, & Gray, 2016).

Gray, 2015; Cheng et al., 2013; Kayyal, Pochedly, McCarthy, & Russell, 2015; Royzman, Atanasov, Landy, Parks, & Gepty, 2014; Schein, Ritter, & Gray, 2016). Instead, the appearance of specificity arises from the lack of appropriate control conditions and statistical procedures that obscure overlap (e.g., ANCOVA; Cameron, Lindquist, & Gray, 2015)

One additional long-standing claim is that "disgust and discomfort drive moral condemnation which are later cloaked with harm based rationalization" (Haidt & Hersh, 2001, p. 212). This suggests that disgust and not harm is the most proximate predictor of moral judgment—especially for ostensibly harmless "purity" acts (Figure 8). We tested this claim in multiple mediation studies where people rated the disgust, harmfulness, and immorality of these acts. These studies revealed that harm statistically mediated the impact of disgust upon moral judgments of these "harmless" acts—often with complete mediation (Schein, Ritter, & Gray, 2016). In other words, disgusting sexual and religious violations are seen as wrong only when feelings of disgust engender perceptions of harm.

Just as perceived harm separates moralized values from nonmoralized values, so too does it separate moralized disgust from nonmoralized disgust. Cleaning up your baby's diarrhea and giving children tainted blood are both disgusting, but only the latter is harmful and therefore immoral. Of course, as negative affect is one key element of immorality, we acknowledge that emotions (including disgust) can have

a causal impact on immorality when controlling for harm. Indeed, past work has linked dyadic elements to dimensions of moral emotions, with anger and disgust tied to viewing immoral agents, and sympathy and sadness tied to viewing victims (Gray & Wegner, 2011c).

The Weakness of Weak Modularity

We recognize that predictions of distinctness, domain- and emotion-specificity are hallmarks of "strong" modularity and there are weaker versions of modularity (Haidt et al., 2015). Although MFT once claimed to be "near the maximalist side of the [modularity] spectrum" (Graham et al., 2013, p. 99), more recent accounts advocate for only "minimal" modularity (Haidt et al., 2015) and focus on the claim that MFT provides only "a first draft of the moral mind" (Graham et al., 2013, p. 63). However, there are two issues with weak modularity: unfalsifiability and lack of explanation.

Unfalsifiability. A strength of strong modularity is that it provides a clearly testable hypothesis—distinctness—and is therefore falsifiable. Unfortunately, weak modularity lacks this falsifiability. How can one test a mechanism that is only metaphorical and makes no claims to separation? Any pattern of data could be consistent with a watered down version of MFT in which moral foundations are merely guiding ideas rather than psychological mechanisms. Even claims of an innate "first draft of the moral mind" shaped by cultural learning are uninformative, as literally all psychological phenomenon hinge on a combination of nature and nurture.

Especially problematic are directly contradictory claims such as both "similarities and differences . . . between different kinds of moral judgment" (Graham, 2015, p. 872), and that moral foundations are both "little switches in the brain" (Haidt, 2012, p. 123), and "not spots in the brain" (Graham et al., 2013, p. 96)—exactly opposite claims that guarantee MFT can never be wrong, whatever the evidence. We suggest that for MFT to be useful, it must specify falsifiable predictions that go beyond the general principles of innateness, cultural learning, intuitionism, and pluralism, especially because these four claims are also endorsed by dyadic morality.

Lack of explanation. The benefit of strong modularity is that it seeks to explain phenomena through the operation of distinct mechanisms (e.g., taste buds). We acknowledge that the existence of distinct, differentially activated moral mechanisms would help to explain moral differences across cultures—however, there is no clear evidence for their existence. Instead, "foundations" are simply tendencies to moralize specific content areas, which means that MFT no longer *explains* moral differences but only *names* them (Gawronski & Bodenhausen, 2015).

Imagine we defined purity as "a tendency to moralize deviant sexuality." This would mean that the statement "Pat morally condemns bestiality because of purity" is equivalent to saying that "Pat moralizes deviant sexuality because she moralizes deviant sexuality"—a clear tautology. To be sure, there are important differences between those who moralize sexuality and those who don't—including religion and conservativism—but simply naming them "purity" obscures any deeper explanation and impedes future research. We suggest that the field should move past these tautological labels and explore the fundamental cognitive and cultural determinants of moral differences.

What Are Moral Foundations?

Although we have criticized much research on modularity and MFT, we suggest that there is a place for moral foundations in the field. However, we stress that empirical evidence suggests that MFT does not describe "foundations"—deep and reified causal mechanisms—but instead genres, similar to music genres. These genres can help catalog and describe the diversity of moral content across cultures and contexts, and hence still retain some "pragmatic utility" (Graham et al., 2013). Unfortunately, even these genres may suffer from problems of construct validity. Not only are harm, fairness, authority, loyalty, and purity highly correlated with each other (Schein & Gray, 2015) but they also have questionable divergent validity from scales like Social Dominance Orientation and Right Wing Authoritarianism (Jost, 2012; Kugler, Jost, & Noorbaloochi, 2014).

One issue is that the items used to assess MFT have conflated conservativism and morality from the very beginning (Graham & Haidt, 2012). Claims that only conservatives are concerned with purity are based upon items tapping sex and religion—two issues long linked to conservatism. However, liberals do seem to care about purity-type concerns such as vaccines and the environment (Frimer et al., 2015), and the sanctity of liberal heroes such as Martin Luther King Jr. (Haidt, 2012). Similarly, classifying liberals as individualists and conservatives as collectivists (Haidt, 2007) does not apply to many important issues (Frimer, Tell, & Motyl, 2016) such as gun rights (conservatives are individualistic) and taxation (liberals are collectivist). Thus, it is unclear whether measures such as the MFQ reveal new knowledge about individual differences in morality or simply repackage a subset of well-known political differences with pithy names.

Also problematic is that MFQ (Graham et al., 2011) uses explicit self-perceptions concerning moral judgment (i.e., "to what extent are the following considerations relevant to your thinking?"), and not the moral judgments themselves. Pioneering work by Haidt (2001) suggested that these explicit self-perceptions do not reliably represent moral cognition, and other work suggests that these responses are subject to exaggeration (Graham, Nosek, & Haidt, 2012).

MFT scenarios tapping moral judgment also confound moral content with weirdness and severity (Gray & Keeney, 2015b), undermining their utility as general-purpose moral questionnaires.

In sum, although the MFQ may be easy to add to studies, researchers would be wise to consider MFT's theoretical and methodological limitations before using its conceptual framework and associated measures—especially when other taxonomies are available (e.g., Janoff-Bulman & Carnes, 2013; Schwartz, 1999).

Is There an Alternative to Modularity? Yes: Constructionism

In contrast to the modular account of morality, dyadic morality denies the existence of natural moral kinds and isomorphic mechanisms—specific cognitive "switches," one for each kind of moral content (Cameron, Lindquist, & Gray, 2015). Instead, TDM suggests that moral content represents varieties of perceived harm. Such variations may be descriptively different—and usefully taxonomized—but are not ontologically distinct because they are made from a common set of core ingredients. Rather than modularity, dyadic morality is inspired by a very different model of the mind: constructionism (L. F. Barrett, 2013).

Constructionism suggests that psychological phenomena emerge from the combination of more basic ingredients, rather than through the operation of distinct mechanisms. As an analogy, consider baked goods (L. F. Barrett, 2009). Croissants, scones, and pancakes each taste and look different, but they all involve various combinations of the same basic ingredients (e.g., flour, sugar, baking powder, salt). In morality, constructionism means that different moralized concerns (e.g., loyalty, purity) consist of various combinations of norms, affect, and perceived harm.

For the sake of argument, imagine that—across cultures—there were five different varieties of agents (e.g., gods, adults, groups), 20 different varieties of acts (e.g., hitting, insulting, defiling), and 15 different varieties of vulnerable patient (e.g., children, adults, animals, souls, social order). When multiplied together, this number would give the possibility for 1,500 varieties of moral judgment—and we haven't even yet considered the nuances of norms.[4] In other words, a synthetic, constructionist understanding of morality offers a combinatorial explosion cultural diversity.

The constructionist roots of TDM rebuffs criticisms that we wish to replace four, five, or six modules with a single neurally instantiated module of harm (Sinnott-Armstrong, 2016). Instead we argue for *emergence* in which a small subset of lower level ingredients can combine to yield diversity at higher levels of analyses. Emergence is more consistent with modern understandings of the evolution of the human mind than is modularity, because emergence is more neurally efficient (L. F. Barrett & Russell, 2014). Thousands of varieties of human mental experiences need not require thousands
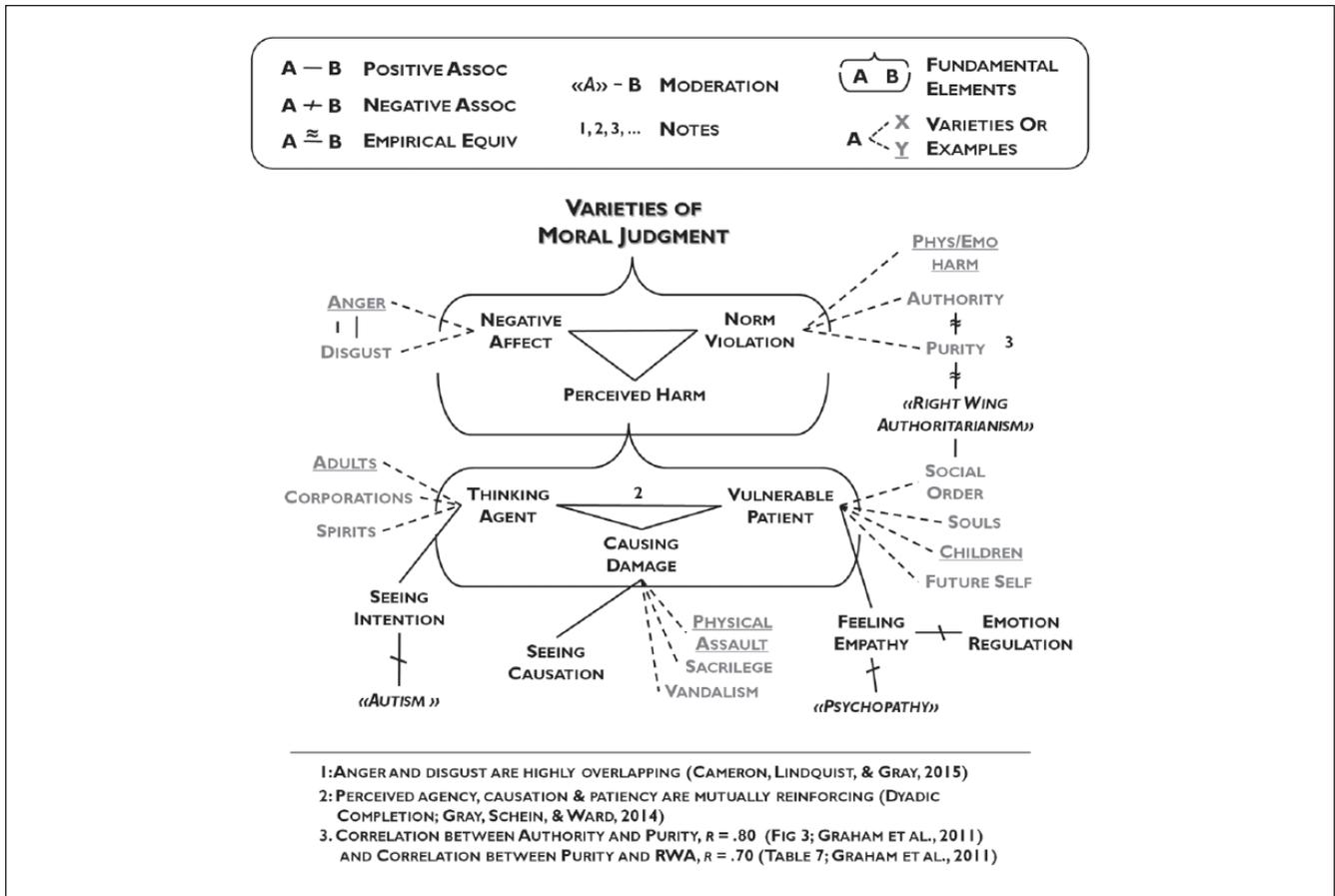
Figure 9. A "Theory Map" of the constructionist Theory of Dyadic Morality (see Gray, in press). Moral judgment emerges from combining varieties of three basic ingredients: Norms, affect, and harm. Harm itself is also constructed of an agent, patient, and a causal damaging act. Each of these features has numerous varieties, allowing for moral diversity. (For more on theory maps, see www. theorymaps.org).

of separate little mental switches, but instead can be made with a small set of domain-general neural networks.

A geometric analogy helps to capture the difference between modularity and constructionism. Modularity is represented by arrows that go from one moral module to one moral concern. In contrast to the arrow, constructionism is illustrated by overlapping circles (think Venn diagrams) in which mental states emerge from the combination of other elements (Gray, Schein, & Cameron, in press). In this case, morality emerges from the overlap of norms, affect, and harm, with harm itself emerging from the overlap of agents, patients, and causation.

This emergence is illustrated in Figure 9 through a "theory map" of dyadic morality (Gray, in press). Elements within the } symbol combine (i.e., overlap) to form the construct to which they point: norms, affect, and harm make immorality; agents, patients, and causation make harm. Varieties of each element are highlighted (e.g., anger and disgust are both forms of negative affect), with the most prototypical variety of each underlined (e.g., children are the most prototypical moral patient). Associations between these elements and various other constructs are also illustrated to situate dyadic morality within the broader framework of moral psychology.

Implications

Beyond suggesting a new perspective on the moral mind, dyadic morality has specific implications for understanding moral character, political discourse, "moral" harms, and cross-cultural dialogue.

Character

Moral psychology has long focused on the immorality of action, but many argue for importance of global, person-centric moral evaluations—judgments of moral character (Pizarro & Tannenbaum, 2011). At first glance, moral character research might seem to contradict dyadic morality (Pizarro et al., 2012), such as when people assign worse moral character to a cat-beater than a woman-beater, despite the fact that the *act* of woman-beating is worse (Tannenbaum,

Uhlmann, & Diermeier, 2011). However, dyadic morality predicts that moral character is ultimately about someone's general capacity for harmful norm violations, and as the authors of this study suggest, cat-beating suggests a more disturbing ability for these actions in general. Consistent with this idea, participants saw the cat-beater as less likely to "help the homeless" and more likely to "enjoy the suffering of others" (Tannenbaum et al., 2011)—both variables related to harm.

Other studies revealed that people morally condemn the character of those who commit "objectively" harmless moral acts (e.g., bizarre masturbation), but again, these acts suggest someone who will violate other harm-based norms (Uhlmann & Zhu, 2013). Consider a man who masturbates with a dead chicken. Although this act is "harmless," parents likely would not want him babysitting their daughter for fear of the harm he might do to her. Consistent with this logic—and with dyadic morality—research finds a cognitive asymmetry between harm and "purity," in which purity violations are tied to potential harm, but not vice versa (Chakroff, 2015).

Dyadic morality also provides a nuanced perspective on moral character, suggesting that character is not only about good versus evil but also about moral agents versus moral patients. Just as Hollywood actors are typecast into enduring Hollywood roles, so too are moral actors typecast as *either* agents (those who do moral acts) *or* patients (those who receive moral acts (Gray & Wegner, 2009). *Moral typecasting* suggests that, in some ways, the psychological distance between hero and villain is relatively small because both are moral agents—explaining why it is easy for heroes to fall from grace (Weeks, 2012). Conversely, the distance between victimhood and villainy/heroism is relatively far, explaining why we assign less blame to victims than to heroes when they perpetrate wrongs (Gray & Wegner, 2011c)—at least when their victimhood does not threaten our self-concept (Gray & Wegner, 2010b).

Typecasting is consistent with classic work in attribution in which ascriptions of responsibility and control are inversely related to ascriptions of victimhood (Weiner, 1995). In the heyday of the AIDS epidemic, research revealed that sympathy for HIV-positive men hinged on whether the sufferer was seen as a helpless, vulnerable sick man (i.e., a patient), or a sinful homosexual responsible for his own misfortune (i.e., an agent; Weiner, 1993). Competing perceptions of agents and patients also explains why heroes and villains are viewed as insensitive to pain: The more you do good or evil, the less you are seen as vulnerable to receiving it.

Political Disagreement—Harm as a Moral Lingua Franca

Dyadic morality suggests that liberals and conservatives do not have fundamentally different moral minds, but instead see harm differently (Schein & Gray, 2015). Consistent with this idea, studies reveal that both liberals and conservatives use considerations of harm to decide which of two acts is more immoral (Schein & Gray, 2015; acts adapted from Graham & Haidt, 2012). Participants were asked to indicate whether violations of loyalty (e.g., burning your country's flag in private when no one else sees you) or fairness (e.g., signing a secret-but-binding pledge to only hire people of your race in your company) were more immoral. Although conservatives saw loyalty violations as more wrong, these comparative moral judgments were underlain by differences in perceived harm (i.e., conservatives also saw loyalty violations as more harmful; Schein & Gray, 2015).

Links between immorality and harm suggest that moral gaps can be reduced by decreasing differences in perceived harm. Consider abortion. Both liberals and conservatives agree that killing mindless cells (e.g., E. coli bacteria) is permissible, and that killing babies is wrong. Where they disagree is whether a 10-week-old fetus is better understood as an insensate mass of cells or a vulnerable baby—a question that may be answered by increased work in fetal medicine. Of course, minds are ultimately a matter of perception, but perceptions can be changed. As late as 1971, evangelical Christians believed that life begins at birth (distinguishing themselves from Catholics), but now believe that life begins at conception (Wegner & Gray, 2016).

Even if perceptions of harm are hard to change, simply acknowledging the legitimacy of perceived harm can help soothe the rancor of moral debates. When it comes to gay marriage and feminism, liberals often believe that conservatives are inventing harm simply to antagonize them. Conversely, conservatives often believe that liberals have a wanton disregard for the safety of children and the future of America (Bryant, 1977; Peters, 2016). By understanding that we all share fundamentally the same moral mind, it may help us humanize our moral opponents. For example, after we wrote an op-ed about how many conservatives see gay marriage as causing legitimate suffering (Gray & Schein, 2015), we received an email from a Baptist pastor who warmly thanked us for recognizing his harm-based position. Simply acknowledging that others perceive harm can foster respect, even when agreement is difficult to reach.

"Moral" Harms

Cultures universally moralize harm, but some harms seem to be encouraged, such as killing enemies in war (Rai & Fiske, 2012) or crushing the opposition in football games (Guttmann, 2004). How does dyadic morality square with such "virtuous violence" (A. P. Fiske & Rai, 2014)? Local norms about what is common or permissible certainly play an important role in these acts, but people may also not perceive substantial harm in them. By dementalizing enemies in war and sports, people fail to see them as capable of suffering and therefore as undeserving of moral concern, an idea supported by substantial research (Goff et al., 2008; Haslam, 2006). Another explanation for these encouraged harms is

that they are actually seen as being helpful overall. By killing an enemy in a foreign country, a soldier may safeguard millions of American lives including his friends and family. By tackling an opposing player in football, a player may be building character or inspiring children to exercise.

Even in cases where we seem to acknowledge the pain of others, it is only unjust if we empathize with it. Empathic reactions may be innate (Preston & de Waal, 2001), but they can be dampened through emotion regulation (Cameron & Payne, 2011), dehumanization (Cameron, Harris, & Payne, 2015), and rationalization (Bandura et al., 1996). Even still, such flickers of empathy can survive, such as when soldiers have difficulty shooting at the opposition (Grossman, 2009), and when individuals defy cultural narratives to help enemies in need (Bilewicz & Jaworska, 2013).

Moral typecasting may also help to explain virtuous violence (Gray & Wegner, 2009). Cultures do not delight at the suffering of all people, but only at a very small group: the evil. By inflicting suffering upon the villainous, they are transformed from agents of evil into mere patients, a process which serves to nullify their evilness.

Cross-Cultural Differences

Although critics have claimed that dyadic morality denies cross-cultural moral diversity (Graham et al., 2013), this characterization is wrong. While TDM may not taxonomize the specific content of cultural variability, it celebrates moral diversity via diversity in values, which are moralized when they are tied to harm. Among all predictions of dyadic morality, this cross-cultural hypothesis is the least tested; however, some cross-cultural data are suggestive.

Classic anthropological accounts from Shweder reveal that people link different kinds of suffering to different kinds of immorality (Shweder, Much, Mahapatra, & Park, 1997). In cultures lacking scientific explanations of suffering, these harmful violations of natural order provide a compelling "moral causal ontology" (Shweder et al., 1997, p. 422), potentially explaining the origin of sanctity concerns. A more recent study finds that cultures moralize chastity and other sanctity-related concerns when the prevalence of sexually transmitted diseases rises (Tybur et al., 2013; van Leeuwen, Park, Koenig, & Graham, 2012), consistent with the causal role of harm.

However, one study does seem to argue against the cross-cultural role of harm by revealing a dissociation between harm and immorality in explicit judgments of Chinese participants (Buchtel et al., 2015). This sample of Chinese participants viewed morality as more about civility than about harm, rating it more immoral to not give up your bus seat to an elderly person than to murder that elderly person. There are several open questions with this study. First, morality is a code that govern behaviors, and given that bus-seat keeping is more prevalent than murder, we must wonder at this conception of "morality." Second, even Americans have a sense

of "politeness," and we wonder if the authors merely measured this construction in Chinese culture. Third, dyadic morality also acknowledges a powerful role of norms in morality, and norms certainly vary across cultures.

As intention is an important part of harm, TDM suggests that concerns about the intention behind immorality should be prevalent cross-culturally. Consistent with this idea, intentionality perception emerges early in development in both industrialized cultures and in rural, small-scale, traditional societies (H. C. Barrett et al., 2013). We note that in one large-scale survey, two traditional societies did not emphasize intentionality in moral judgments (H. C. Barrett et al., 2016), a phenomenon explained by cultural learning. For example, one of these societies was the Fijian Yasawan, who endorse Opacity of Mind norms, the belief that other minds are unknowable and should not be discussed (J. Robbins & Rumsey, 2008). Nevertheless, experimentally focusing Yasawan adults on the thoughts behind actions led to greater moral differentiation of failed attempts versus accidents, consistent with the role of intention (McNamara, 2016). Furthermore, Yasawan children differentiate intent in similar fashion as Western children (Hamlin, 2013; McNamara, 2016), suggesting an innate moral sensitivity to intentionality that cultural learning is needed to suppress.

Testing Dyadic Morality

Dyadic morality suggests a new model of moral judgment, and although we have reviewed substantial indirect and direct evidence in its support, additional research is surely needed. Here we provide some recommendations for study design which focus primarily upon the measurement of harm. We predict that—when they are followed—future studies will reveal a robust association between perceived harm and moral judgment. Dyadic morality would be falsified if—when controlling for norms and affect—intuitive perceptions of harm are not causally linked to intuitive perception of immorality.

Although TDM's redefinition of harm makes it more challenging to study, each of its predictions are also falsifiable. The intuitiveness of harm means that harm should be more bound to moral judgment when it is assessed under cognitive load or with implicit measures. This is exactly what studies reveal when assessing judgments of harm and immorality regarding "purity" violations with time pressure and implicit measures (Gray, Schein et al., 2014). The perceptual nature of harm also means that participants' reports of harm should better predict moral judgment than the "objective" harm designed by researchers within scenarios.

A continuum of harm means that ratings of harm should span a range rather than being bimodal, as revealed by previous research (Cushman et al., 2006; Schein & Gray, 2015). The synthetic nature of harm means that ratings of harm—and immorality—should be sensitive to manipulations of intention, causation, and suffering, consistent with empirical

evidence (Ames & Fiske, 2013; Cushman, 2015; Schein & Gray, 2015). Overall, testing the claims of the "new" intuitive harm should be no more difficult than testing the claims of the "new" intuitive moral judgment, as social psychology as been doing since 2001 (Greene et al., 2001; Haidt, 2001).

Study Design

Intuitive perceived continuum. Previous research assumed that acts can be objectively "harmless" (Haidt & Hersh, 2001); however, TDM suggests that both harm and morality are intuitively perceived. Accordingly, participants should be allowed to indicate how harmful or immoral an act *seems* to them. Of course, while responses to explicit questions can help shed light on the structure of our moral cognition, they might also tap effortful reasoning, so ideally perceptions should be accessed with implicit measures or under time pressure to better tap intuitions. Given that both harm and morality exist along a continuum, researchers should ask *how* harmful or immoral an act seems and not just provide binary options (Piazza & Sousa, 2016).

Synthetic. Harm is not merely suffering, but the dyadic combination of an intentional agent, causal damage, and vulnerable patient. Questions should therefore assess both global perceptions of harm, and the specific elements of the dyad—along with relevant perceptions of mind such as agent's intention and a patient's vulnerability, both in terms of general capacity (e.g., how capable is *x* of suffering or intention?), and its specific manifestation in an act (e.g., how much did *x* intend or suffer from this act?).

Controls. To test the specific role of harm, and to separate "moral wrongness" from "general wrongness" studies should also assess the role of general negativity (i.e., unpleasantness) and also weirdness and typicality, which are tied to norms (Gray & Keeney, 2015b). For more discussion on appropriate controls, see Cameron, Lindquist, and Gray (2015).

Capturing the full continuum of immorality and harm. Scenarios used to tap moral judgment should assess the full continuum of immorality and harm by including scenarios that are canonically immoral (e.g., murder), canonically not immoral (e.g., jogging), and all points between. Without this full range, erroneous conclusions could be reached. For example, most studies testing the disgust–immorality link use only morally ambiguous and disgusting scenarios such as eating a dead dog or buying sexually explicit music (Haidt, 2001; Horberg et al., 2009). When a broader set of scenarios is used (e.g., murder and cleaning up poop), the unique impact of disgust disappears and the role of harm is clearly apparent (Schein & Gray, 2015). Ideally, scenarios should also capture variation along many dimensions, including act severity, weirdness, intentionality, general negativity, and the specific value being violated.

Naturalistic scenarios. Morality evolved to deal with challenges typical in ancestral societies, and cultural learning has evolved to focus upon challenges typical in today's societies. To assess moral cognition, scenarios should tap these everyday challenges rather than bizarre cases invented by researchers. For example, to assess morality related to purity, we have developed a set of stimuli reflecting naturalistic violations (e.g., prostitution, adultery; Gray & Keeney, 2015b).

Areas of Investigation

Tests of dyadic morality may be especially useful when they interface with four emerging areas in moral psychology.

Moralization. Anthropological and historical research has shown how moral judgments evolve over time and place (e.g., Laqueur, 2004). Dyadic morality predicts that moralization is driven by changes in perceived harm, whether because of systemic forces or random variation (Schein & Gray, 2017). The coherence processes of dyadic morality also predicts that it will be easier for an act to enter the moral sphere than exit it. Once a dyadic loop is initiated and an act is categorized as immoral, rationally eliminating one type of harm might not be sufficient, as dyadic completion will likely lead to harm being perceived elsewhere (for expanded predictions of moralization and demoralization, see Schein & Gray, 2016; 2017).

Religion. Religion and morality are intricately connected, with religious doctrine informing moral judgments, and binding individuals into communities (Graham & Haidt, 2010; Nucci & Turiel, 1993). Dyadic morality predicts that religious moral doctrine is shaped by historical and evolutionary threats (Douglas, 1966), which then shapes proximal perceptions of harm, consistent with the dyad loop.

Taxonomies. More empirical research is also needed to help bridge TDM with work categorizing moral diversity (Haidt et al., 1993; Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011). For example, one could integrate TDM with the Relationship Regulation Model (Rai & Fiske, 2011), asking how different types of relationship structures transform perceptions of harm. Theoretical models can also help organize types of harms to different patients (e.g., the self, others, society), as in the MMM (Janoff-Bulman & Carnes, 2013).

Individual differences. TDM predicts that individual differences in threat sensitivity (i.e., perceived potential harm) should correspond to moral condemnation. This prediction is consistent with past findings showing that individual differences in threat sensitivity underlies political differences (Jost et al., 2007), and the differential moral condemnation of both disgust (Schein, Ritter, & Gray, 2016) and diverse violations (van Leeuwen & Park, 2009). Future research should also explore systematic variations in where exactly people see harm.

Conclusion

Viable theories of moral cognition should address several key questions: What constitutes moral judgments and differentiates them from other types of judgments? What accounts for moral pluralism? How are moral judgments made? How do acts enter the moral realm, and how are they entrenched?

Dyadic morality provides testable answers to each of these questions—answers which highlight the importance of harm. This is not the reasoned, objective, and binary harm of old, but instead an intuitively perceived continuum, which is instantiated as a fuzzy cognitive template. This cognitive template is rooted in evolutionary concerns about proximate harm, but can be modified by cultural assumptions about more distal harms.

Consistent with traditional harm-centric accounts, TDM suggests that harm separates "conventional" negative norm violations from violations of morality (Turiel, 1983). However, consistent with modern moral pluralism, TDM allows for cultural diversity through the flexibility of harm and a variety of values. By allowing for diversity within a single synthetic template, dyadic morality provides parsimony and pluralism.

There are two quotes which best summarize the aspirations of dyadic morality. The first is from moral anthropologist Rick Shweder (2012), who advocates for "universality without the uniformity" (p. 88). We suggest that a flexible harm-based template allows for a universal understanding of morality in tandem with rich cultural diversity. The second quote is from Felix Mendelssohn, who suggested that "the essence of the beautiful is unity in variety." The essence of morality is also unity in variety—cognitive unity in the variety of perceived harm.

Notes

1. More than 80% of people intuitively judge acts like these as immoral (Hester, Payne, & Gray, 2017)
2. In self-defense (e.g., shooting someone who is stabbing you), it is unclear which person is the agent and patient, causing ambiguity in moral judgment.
3. An interesting tidbit on the potential randomness of moral change—one historian suggests that the pamphlet's anonymous writer "invented a new disease" as a way to earn an income selling snake oil (Laqueur, 2004, p. 14).
4. Of course, these numbers may not be fully independent, as some acts (e.g., defiling) better align with certain patients (e.g., souls), but the general point of combinatorial explosion still holds.

References

Adam, B. D. (2003). The Defense of Marriage Act and American exceptionalism: The gay marriage panic in the United States. *Journal of the History of Sexuality*, *12*, 259-276. doi:10.1353/sex.2003.0074

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

Alicke, M. D. (2012). Self-injuries, harmless wrongdoing, and morality. *Psychological Inquiry*, *23*, 125-128. doi:10.1080/1047840X.2012.666720

Ames, D. L., & Fiske, S. T. (2013). Intentional harms are worse, even when they're not. *Psychological Science*, *24*, 1755-1762.

Anderson, R. (2013, March 18). *Marriage matters: Consequences of redefining marriage*. The Heritage Foundation. Retrieved from http://www.heritage.org/research/reports/2013/03/why-marriage-matters-consequences-of-redefining-marriage

Aristotle. (BC350). *Nicomachean ethics* (W. D. Ross, Trans.). New York, NY: World Library Classics.

Asch, S. E. (1952). *Social psychology* (1st ed.). Englewood Cliffs, NJ: Prentice Hall.

The Associated Press. (2016, September 10). 2nd suspect in "Slender Man" stabbing enters insanity plea. *NBC News*. Retrieved from http://www.nbcnews.com/storyline/slender-man-stabbing/2nd-suspect-slender-man-stabbing-enters-insanity-plea-n645976

Atran, S., & Ginges, J. (2012). Religious and sacred imperatives in human conflict. *Science*, *336*, 855-857.

Avramides, A. (2001). *Other minds*. London, UK: Routledge.

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review*, *80*, 1095-1111.

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.

Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, *17*, 1-10.

Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes*, *94*(2), 74-85.

Baron-Cohen, S. (2011). *The science of evil: On empathy and the origins of cruelty*. New York, NY: Basic Books.

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., Gurven, M., . . . Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, *113*, 4688-4693. doi:10.1073/pnas.1522070113

Barrett, H. C., Broesch, T., Scott, R. M., He, Z., Baillargeon, R., Wu, D., . . . Laurence, S. (2013). Early false-belief understanding in traditional non-Western societies. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*, 20122654. doi:10.1098/rspb.2012.2654

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior*, *26*, 313-331. doi:10.1016/j.evolhumbehav.2004.08.015

Barrett, L. F. (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition & Emotion*, *23*, 1284-1306. doi:10.1080/02699930902985894

Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review*, *5*, 379-389. doi:10.1177/1754073913489753

Barrett, L. F., & Russell, J. A. (Eds.). (2014). *The psychological construction of emotion* (1st ed.). New York, NY: The Guilford Press.

Bastian, B., & Crimston, D. (2016). The moral psychology of resource use. In J. P. Forgas, L. Jussim, & P. A. M. Van Lange (Eds.), *The social psychology of morality* (p. 274-288). London, UK: Routledge.

Bastian, B., Loughnan, S., Haslam, N., & Radke, H. R. M. (2012). Don't mind meat? The denial of mind to animals used for human consumption. *Personality and Social Psychology Bulletin*, *38*, 247-256. doi:10.1177/0146167211424291

Benatar, D., & Benatar, M. (2001). A pain in the fetus: Toward ending confusion about fetal pain. *Bioethics*, *15*(1), 57-76. doi:10.1111/1467-8519.00212

Bentham, J. (1970). An introduction to the principles of morals and legislation. In J. H. Burns & H. L. A. Hart (Eds.), *Collected works of Jeremy Bentham*. London, UK: Athlone. (Original work published 1780)

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge, UK: Cambridge University Press.

Bicchieri, C., & Muldoon, R. (2014). Social norms. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/spr2014/entries/social-norms/

Bilewicz, M., & Jaworska, M. (2013). Reconciliation through the righteous: The narratives of heroic helpers as a fulfillment of emotional needs in Polish−Jewish intergroup contact. *Journal of Social Issues*, *69*, 162-179. doi:10.1111/josi.12009

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

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

Brandt, M. J., Wisneski, D. C., & Skitka, L. J. (2015). Moralization and the 2012 US Presidential Election Campaign. *Journal of Social and Political Psychology*, *3*, 211-237.

Bratanova, B., Loughnan, S., & Bastian, B. (2011). The effect of categorization as food on the perceived moral standing of animals. *Appetite*, *57*, 193-196. doi:10.1016/j.appet.2011.04.020

Brauer, M., & Chaurand, N. (2010). Descriptive norms, prescriptive norms, and social control: An intercultural comparison of people's reactions to uncivil behaviors. *European Journal of Social Psychology*, *40*, 490-499. doi:10.1002/ejsp.640

Brauer, M., & Chekroun, P. (2005). The relationship between perceived violation of social norms and social control: Situational factors influencing the reaction to deviance. *Journal of Applied Social Psychology*, *35*, 1519-1539.

Brown, E. N. (2014, December 2). U.K. censors online porn; spanking, rough sex, other assorted kink banned. Hit & Run. Retrieved from http://reason.com/blog/2014/12/02/uk-bans-fetish-porn

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.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., & Bond, M. H. (2015). Immorality East and West are immoral behaviors especially harmful, or especially uncivilized? *Personality and Social Psychology Bulletin*, *41*, 1382-1394. doi:10.1177/0146167215595606

Cameron, C. D., Harris, L. T., & Payne, B. K. (2015). The emotional cost of humanity anticipated exhaustion motivates dehumanization of stigmatized targets. *Social Psychological and Personality Science*, *7*, 105-112. doi:10.1177/1948550615604453

Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review*, *19*, 371-394. doi:10.1177/1088868314566683

Cameron, C. D., & Payne, B. K. (2011). Escaping affect: How motivated emotion regulation creates insensitivity to mass suffering. *Journal of Personality and Social Psychology*, *100*(1), 1-15.

Carcaño v. McCrory, No. 1:16-cv-236 (United States District Court for the Middle District of North Carolina 2016). Retrieved from https://www.aclu.org/sites/default/files/field_document/dkt_1_-_carcano_v._mccrory_complaint.pdf

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

Chakroff, A. (2015). *Discovering structure in the moral domain* (Doctoral dissertation). Harvard University, Cambridge, MA.

Cheng, J. S., Ottati, V. C., & Price, E. D. (2013). The arousal model of moral condemnation. *Journal of Experimental Social Psychology*, *49*, 1012-1018. doi:10.1016/j.jesp.2013.06.006

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. *24*, pp. 201-234). San Diego, CA: Academic Press.

Clark, C. J., Chen, E. E., & Ditto, P. H. (2015). Moral coherence processes: Constructing culpability and consequences. *Current Opinion in Psychology*, *6*, 123-128.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*, *47*(4), 1-21. doi:10.3758/s13428-014-0551-2

Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, *81*, 697-710. doi:10.1037/0022-3514.81.4.697

Comer, M. (2012, April 24). Anti-gay N.C. pastor: Amendment One defeat will cause "nuclear holocaust." InterstateQ.com [Blog]. Retrieved from http://interstateq.com/archives/5216/

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation

approach. *Journal of Personality and Social Psychology*, *104*, 216-235.

Cooley, E., Payne, B. K., Cipolli, W., Cameron, D, Berger, A., & Gray, K. (2017). The paradox of group mind: "people in a group" have more mind than "a group of people." *Journal of Experimental Psychology: General*, *146*(5), 691-699.

Corneille, O., Hugenberg, K., & Potter, T. (2007). Applying the attractor field model to social cognition: Perceptual discrimination is facilitated, but memory is impaired for faces displaying evaluatively congruent expressions. *Journal of Personality and Social Psychology*, *93*, 335-352.

Crimston, D., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). *Moral expansiveness: Examining variability in the extension of the moral world*. Retrieved from http://psycnet.apa.org/psycinfo/2016-01230-001/

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. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology*, *6*, 97-103. doi:10.1016/j.copsyc.2015.06.003

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

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

Day, N. (2013, June 21). How to address the masturbating child. *The Atlantic*. Retrieved from http://www.theatlantic.com/health/archive/2013/06/how-to-address-the-masturbating-child/277026/

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, *59*, 279-300. doi:10.1146/annurev.psych.59.103006.093625

Decety, J., & Cacioppo, S. (2012). The speed of morality: A high-density electrical neuroimaging study. *Journal of Neurophysiology*, *108*, 3068-3072. doi:10.1152/jn.00473.2012

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 71-100.

Definition of Harm. (2016). In *Merriam-Webster*. Retrieved from http://www.merriam-webster.com/dictionary/harm

Derbyshire, D. (2010, March). *Scientists discover moral compass in the brain which can be controlled by magnets*. Retrieved from http://www.dailymail.co.uk/sciencetech/article-1262074/Scientists-discover-moral-compass-brain-controlled-magnets.html

Descartes, R. (1641). *The philosophical writings of Rene Descartes* (J. Cottingham, R. Stoothoff, & D. Murdoch, Trans.). Cambridge, UK: Cambridge University Press.

DeScioli, P., Bruening, R., & Kurzban, R. (2011). The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior*, *32*, 204-215.

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

Dijker, A. J. M. (2010). Perceived vulnerability as a common basis of moral emotions. *British Journal of Social Psychology*, *49*, 415-423.

Ditto, P. H., & Liu, B. (2011). Deontological dissonance and the consequentialist crutch. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 51-70). Washington, DC: American Psychological Association.

Ditto, P. H., Liu, B., & Wojcik, S. P. (2012). Is anything sacred anymore. *Psychological Inquiry*, *23*, 155-161.

Douglas, P. M. (1966). Purity and danger: An analysis of concepts of pollution and taboo. New York, NY: Routledge.

Dred Scott v. Sandford, No. 38 (United States Supreme Court 1856).

Dworkin, A. (1981). *Pornography: Men possessing women* (First Perigee Printing ed.). London, England: The Women's Press.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, *87*, 876-893. doi:10.1037/0022-3514.87.6.876

Effron, D. A., & Miller, D. T. (2012). How the moralization of issues grants social legitimacy to act on one's attitudes. *Personality and Social Psychology Bulletin*, *38*, 690-701. doi:10.1177/0146167211435982

Eibach, R. P., Libby, L. K., & Ehrlinger, J. (2009). Priming family values: How being a parent affects moral evaluations of harmless but offensive acts. *Journal of Experimental Social Psychology*, *45*, 1160-1163. doi:10.1016/j.jesp.2009.06.017

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*(1), 91-119.

Eisenberg, N., Spinrad, T. L., & Sadovsky, A. (2006). Empathy-related responding in children. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 517-549). Mahwah, NJ: Lawrence Erlbaum.

Ellis, J. W. (2003). Mental retardation and the death penalty: A guide to state legislative issues. *Mental & Physical Disability Law Reporter*, *27*, 11-24.

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

Epperson, A. R. (1985). *The unseen hand: An introduction to the conspiratorial view of history* (Later Printing ed.). Tucson, AZ: Publius Press.

Eshleman, A. (2014). Moral responsibility. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/sum2014/entries/moral-responsibility/

Feigenson, N., & Park, J. (2006). Emotions and attributions of legal responsibility and blame: A research review. *Law and Human Behavior*, *30*, 143-161.

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*, 178-194. doi:10.1016/0022-1031(85)90014-9

Fincham, F. D., & Shultz, T. R. (1981). Intervening causation and the mitigation of responsibility for harm. *British Journal of Social Psychology*, *20*, 113-120. doi:10.1111/j.2044-8309.1981.tb00483.x

Fiske, A. P. (2000). Complementarity theory: Why human social capacities evolved to require cultural complements. *Personality and Social Psychology Review*, *4*, 76-94. doi:10.1207/S15327957PSPR0401_7

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

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

Fiske, S. T., & Taylor, S. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill.

Fletcher, G. P. (1998). *Basic concepts of criminal law*. Oxford University Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=eCa55NQ8UpIC&oi=fnd&pg=PR9&dq=Basic+Concepts+of+Criminal+Law,&ots=O8MGoDx7Wl&sig=P0lh_fsl-vLnhhDWZe7E0BkUIYs

Fodor, J. A. (1985). Precis of the modularity of mind. *Behavioral and Brain Sciences*, *8*(1), 1-5.

Foer, J. S. (2010). *Eating animals*. New York, NY: Back Bay Books.

Frimer, J. A., Gaucher, D., & Schaefer, N. K. (2014). Political conservatives' affinity for obedience to authority is loyal, not blind. *Personality and Social Psychology Bulletin*, *40*, 1205-1214.

Frimer, J. A., Tell, C. E., & Haidt, J. (2015). Liberals condemn sacrilege too: The harmless desecration of Cerro Torre. *Social Psychological and Personality Science*, *6*, 878-886. doi:10.1177/1948550615597974

Frimer, J. A., Tell, C. E., & Motyl, M. (2016). Sacralizing liberals and fair-minded conservatives: Ideological symmetry in the moral motives in the culture war. *Analyses of Social Issues and Public Policy*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/asap.12127/full

Fullinwider, R. K. (1994). Chastity, morality, and the schools. *Philosophical Dimensions of Public Policy*, *13*, 39.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly*, *46*, 23-30.

Gall, T. L. (2004). Relationship with God and the quality of life of prostate cancer survivors. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*, *13*, 1357-1368.

Gawronski, B., & Beer, J. S. (2016). What makes moral dilemma judgments "utilitarian" or "deontological"? *Social Neuroscience*. Advance online publication. doi:10.1080/17470919.2016.1248787

Gawronski, B., & Bodenhausen, G. V. (2015). Theory and explanation in social psychology. In *Theory and explanation in social psychology*. Guilford Press. Retrieved from http://www.guilford.com/excerpts/gawronski3.pdf

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., . . . Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*, 1100-1104. doi:10.1126/science.1197754

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115-140. doi:10.1146/annurev.psych.59.103006.093659

Gert, B. (2015). The definition of morality. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/fall2015/entries/morality-definition/

Ginges, J., Atran, S., Medin, D., & Shikaki, K. (2007). Sacred bounds on rational resolution of violent political conflict. *Proceedings of the National Academy of Sciences*, *104*, 7357-7360. doi:10.1073/pnas.0701768104

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

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition*, *106*, 1339-1366. doi:10.1016/j.cognition.2007.06.007

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology*, *48*, 250-256. doi:10.1016/j.jesp.2011.08.006

Govrin, A. (2014). The ABC of moral development: An attachment approach to moral judgment. *Frontiers in Psychology, 5*, Article 6. doi:10.3389/fpsyg.2014.00006

Graham, J. (2015). Explaining Away Differences in Moral Judgment: Comment on Gray and Keeney (2015). *Social Psychological and Personality Science*, *6*(8), 869-873. https://doi.org/10.1177/1948550615592242

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. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. Shaver & M. Mikulincer (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 11-31). New York, NY: APA Books.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55-130. doi:10.1016/B978-0-12-407236-7.00002-4

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. (2012). The moral stereotypes of liberals and conservatives: Exaggeration of differences across the political spectrum. *PLoS ONE*, *7*, e50092.

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, K. (2014). Harm concerns predict moral judgments of suicide: Comment on Rottman, Kelemen and Young (2014). *Cognition*, *133*, 329-331. doi:10.1016/j.cognition.2014.06.007

Gray, K. (in press). How to map theory: Reliable methods are fruitless without rigorous theory. *Perspectives on Psychological Science*.

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., & Keeney, J. E. (2015a). Disconfirming moral foundations theory on its own terms: Reply to Graham (2015). *Social Psychological and Personality Science*, *6*, 874-877.

Gray, K., & Keeney, J. E. (2015b). Impure, or just weird? Scenario sampling bias raises questions about the foundation of moral cognition. *Social Psychological and Personality Science*, *6*, 859-868.

Gray, K., Rand, D. G., Eyal, E., Lewis, K., Hershman, S., & Norton, M. I. (2014). The emergence of "us and them" in 80 lines of

code: Modeling group genesis in homogenous populations. *Psychological Science*, *25*, 982-990.

Gray, K., & Schein, C. (2012). Two minds vs. two philosophies: Mind perception defines morality and dissolves the debate between deontology and utilitarianism. *Review of Philosophy and Psychology*, *3*(3), 1-19. doi:10.1007/s13164-012-0112-5

Gray, K., & Schein, C. (2015, January 30). The myth of the harmless wrong. *The New York Times*. Retrieved from http://www.nytimes.com/2015/02/01/opinion/sunday/the-myth-of-the-harmless-wrong.html

Gray, K., & Schein, C. (2016). No absolutism here: Harm predicts moral judgment 30x better than disgust—Commentary on Scott, Inbar & Rozin (2015). *Perspectives on Psychological Science*, *11*, 325-329.

Gray, K., Schein, C., & Cameron, C. D. (in press). How to think about emotions and morality: Circles, not arrows. *Current Opinion in Psychology*.

Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143*, 1600-1615. doi:10.1037/a0036149

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, *23*, 206-215.

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). Morality takes two: Dyadic morality and mind perception. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil*. Washington, DC: APA Press.

Gray, K., & Wegner, D. M. (2011c). 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

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101-124. doi:10.1080/1047840x.2012.651387

Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New York, NY: Penguin.

Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*, 517-523.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144-1154.

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

Grossman, D. (2009). *On killing: The psychological cost of learning to kill in war and society* (Rev. ed.). New York, NY: Back Bay Books.

Gutierrez, R., & Giner-Sorolla, R. (2007). Anger, disgust, and presumption of harm as reactions to taboo-breaking behaviors. *Emotion*, *7*, 853-868. doi:10.1037/1528-3542.7.4.853

Guttmann, A. (2004). *From ritual to record: The nature of modern sports*. New York, NY: Columbia University Press.

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

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

Haidt, J., Graham, J., & Ditto, P. (2015, October 28). The Volkswagen of moral psychology. Retrieved from http://www.spsp.org/blog/dyadic-morality

Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*, 191-221.

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The Innate Mind*, *3*, 367-392.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske & D. Gilbert (Eds.), *Handbook of social psychology* (5th ed., pp. 797-832). Hoboken, NJ: John Wiley.

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

Halevy, N., Kreps, T. A., Weisel, O., & Goldenberg, A. (2015). Morality in intergroup conflict. *Current Opinion in Psychology*, *6*, 10-14. doi:10.1016/j.copsyc.2015.03.006

Hamlin, J. K. (2012). A developmental perspective on the moral dyad. *Psychological Inquiry*, *23*, 166-171. doi:10.1080/1047840X.2012.670101

Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*, 451-474.

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*, 30-39. doi:10.1016/j.cogdev.2010.09.001

Harris, P. L. (2015). *Trusting what you're told: How children learn from others* (Reprint ed.). Cambridge, MA: Belknap Press.

Harrison, S. (2016, March 23). N.C. Gov Pat McCrory signs into law bill restricting LGBT protections. *The Charlotte Observer*. Retrieved from http://www.charlotteobserver.com/news/local/article67845317.html

Hart, H. L. A. (2008). *Punishment and responsibility: Essays in the philosophy of law*. Oxford, UK: Oxford University Press.

Hart, H. L. A., & Honoré, T. (1985). *Causation in the law* (2nd ed.). New York, NY: Oxford University Press.

Hartshorne, J. K. (2014). What is implicit causality? *Language, Cognition and Neuroscience*, *29*, 804-824. doi:10.1080/01690965.2013.796396

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252-264.

Haslam, N. (2012). Morality, mind, and humanness. *Psychological Inquiry*, *23*, 172-174. doi:10.1080/1047840X.2012.655236

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry*, *27*, 1-17.

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

Helion, C., & Pizarro, D. (2015). Beyond dual-processes: The interplay of reason and emotion in moral judgment. In J. Clausen & N. Levy (Eds.), *Handbook of neuroethics* (pp. 109-125). New York, NY: Springer.

Helzer, E. G., & Pizarro, D. (2011). Dirty liberals! Reminders of physical cleanliness influence moral and political attitudes. *Psychological Science*, *22*, 517-522.

Hesse, E., Mikulan, E., Decety, J., Sigman, M., del Carmen Garcia, M., Silva, W., . . . Ibanez, A. (2015). Early detection of intentional harm in the human amygdala. *Brain*, *139*, 54-61.

Hester, N., & Gray, K. (2017). *Dyadic reversal: Implicit causality reveals when victims are seen as perpetrators*. Unpublished manuscript, University of North Carolina, Chapel Hill.

Hester, N., Payne, B. K., & Gray, K. (2017). Emergent morality: The minimal (and meaningless) context sufficient for moral judgment. Unpublished manuscript, University of North Carolina, Chapel Hill.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, *113*, 4296-4301. doi:10.1073/pnas.1516047113

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*, 1340-1343.

Holmes, O. W. (1909). *The common law*. Boston, MA: Little, Brown.

Horberg, E. J., Oveis, C., & Keltner, D. (2011). Emotions as moral amplifiers: An appraisal tendency approach to the influences of distinct emotions upon moral judgment. *Emotion Review*, *3*, 237-244.

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*, 963-976.

H.R. 356 - 109th Congress. (2005). Unborn Child Pain Awareness Act of 2005 (2005 - S. 51). Retrieved from https://www.govtrack.us/congress/bills/109/s51

Huebner, B., Lee, J. J., & Hauser, M. D. (2010). The moral-conventional distinction in mature moral competence. *Journal of Cognition and Culture*, *10*(1), 1-26.

Hume, D. (1740). *A treatise of human nature*. New York, NY: Dover.

Hume, D. (1777). *An enquiry concerning the principles of morals*. La Salle, IL: Open Court.

Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social-functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*, 719-737.

Hyslop, A. (2015). Other minds. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/fall2015/entries/other-minds/

Inbar, Y., Pizarro, D., & Bloom, P. (2012). Disgusting smells cause decreased liking of gay men. *Emotion*, *12*(1), 23-27. doi:10.1037/a0023984

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS ONE*, *7*(8), e42366. doi:10.1371/journal.pone.0042366

James, K. (2010). *Is the "Harry Potter . . . " series truly harmless?* Retrieved from http://www.christiananswers.net/q-eden/harrypotter.html

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape moral motives and group-based moralities. *Personality and Social Psychology Review*, *17*, 219-236. doi:10.1177/1088868313480274

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

Jenni, K., & Loewenstein, G. (1997). Explaining the identifiable victim effect. *Journal of Risk and Uncertainty*, *14*, 235-257.

Johnson, D. J., Wortman, J., Cheung, F., Hein, M., Lucas, R. E., Donnellan, M. B., . . . Narr, R. K. (2016). The effects of disgust on moral judgments: Testing moderators. *Social Psychological and Personality Science*, *7*, 640-647. doi:10.1177/1948550616654211

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Morristown, NJ: General Learning Press.

Jost, J. T. (2012). Left and right, right and wrong. *Science*, *337*, 525-526. doi:10.1126/science.1222565

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

Kant, I. (1780). *The metaphysical elements of ethics* (T. K. Abbott, Trans.). Fairford, UK: Echo Library.

Kayyal, M. H., Pochedly, J., McCarthy, A., & Russell, J. A. (2015). On the limits of the relation of disgust to judgments of immorality. *Frontiers in Psychology*, *6*, Article 951. doi:10.3389/fpsyg.2015.00951

Kellogg, J. H. (1890). *Plain facts for old and young: Embracing the natural history and hygiene of organic life*. Burlington, IA: I. F. Segner. Retrieved from http://archive.org/details/plainfactsforol00kell

Kennedy, J. A. Obergefell et al. v. Hodges, Director, Ohio Department of Health, et al., No.14-556 (Supreme Court of the United States June 26, 2015).

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

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., & Sloan, S. (1948). *Sexual behavior in the human male*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2092220/pdf/brmedj03707-0002.pdf

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

Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The development of self-regulation in the first four years of life. *Child Development*, *72*, 1091-1111.

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, *15*, 694-698.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In T. Mischel (Ed.), *Cognitive development and epistemology* (pp. 151-235). New York, NY: Academic Press.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*, 97-109. doi:10.1037/a0020762

Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another look at Moral Foundations Theory: Do authoritarianism and social dominance orientation explain liberal-conservative differences in "moral" intuitions? *Social Justice Research*, *27*, 413-431.

Lakshminarayanan, V. R., & Santos, L. R. (2008). Capuchin monkeys are sensitive to others' welfare. *Current Biology*, *18*, R999-R1000.

Landy, J. F. (2016). Representations of moral violations: Category members and associated features. *Judgment and Decision Making*, *11*, 496-508.

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, *10*, 518-536. doi:10.1177/1745691615583128

Laqueur, T. W. (2004). *Solitary sex: A cultural history of masturbation* (New ed.). New York, NY: Zone Books.

Layman, G. (2001). *The great divide: Religious and cultural conflict in American party politics*. New York, NY: Columbia University Press.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*, 265-288. doi:10.1016/S0010-0277(87)80006-9

Leslie, A. M., Mallon, R., & DiCorcia, J. A. (2006). Transgressors, victims, and cry babies: Is basic moral judgment spared in autism? *Social Neuroscience*, *1*, 270-283. doi:10.1080/17470910600992197

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist model. *Emotion Review*, *5*, 356-368.

Lindquist, K. A., & Barrett, L. F. (2012). A functional architecture of the human brain: Emerging insights from the science of emotion. *Trends in Cognitive Sciences*, *16*, 533-540.

Lindquist, K. A., Siegel, E. H., Quigley, K. S., & Barrett, L. F. (2013). The hundred-year emotion war: Are emotions natural kinds or psychological constructions? Comment on Lench, Flores, and Bench (2011). *Psychological Bulletin*, *139*, 255-263. doi:10.1037/a0029038

Liu, B., & Ditto, P. H. (2013). What dilemma? Moral evaluation shapes factual belief. *Social Psychological and Personality Science*, *4*, 316-323.

Lukianoff, G., & Haidt, J. (2015, September). The coddling of the American mind. *The Atlantic*. Retrieved from http://www.the-atlantic.com/magazine/archive/2015/09/the-coddling-of-the-american-mind/399356/

Machery, E. (2012). Delineating the moral domain. *Baltic International Yearbook of Cognition, Logic and Communication*, *7*(1), 6.

Malle, B. F. (2006). Intentionality, morality, and their relationship in human judgment. *Journal of Cognition and Culture*, *6*, 87-112.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*, 147-186. doi:10.1080/1047840X.2014.877340

Marmor, A., & Sarch, A. (2015). The nature of law. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/fall2015/entries/law-phil-nature/

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

Martin, J. W., & Cushman, F. (2016a). The adaptive logic of moral luck. In J. Sytsma & W. Buckwalter (Eds.), *The Blackwell companion to experimental philosophy* (pp. 190-202). Chichester, UK: John Wiley & Sons.

Martin, J. W., & Cushman, F. (2016b). Why we forgive what can't be controlled. *Cognition*, *147*, 133-143. doi:10.1016/j.cognition.2015.11.008

Masserman, J. H., Wechkin, S., & Terris, W. (1964). "Altruistic" behavior in rhesus monkeys. *The American Journal of Psychiatry*, *121*, 584-585.

Mayer, B., Merckelbach, H., & Muris, P. (2000). Self-reported automaticity and irrationality in spider phobia. *Psychological Reports*, *87*, 395-405.

McNamara, R. A. (2016, August). *Morality when the mind is opaque: Intent vs. outcome across the lifespan in Yasawa, Fiji*. Vancouver, Canada: The University of British Columbia. Retrieved from https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0307352

Mead, M. (1961). Some anthropological considerations concerning natural law. *Natural Law Forum*, *6*, Article 51.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, *44*, 1469-1481. doi:10.1037/0003-066X.44.12.1469

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

Meigs, A. S. (1984). *Food, sex, and pollution: A New Guinea religion*. New Brunswick, NJ: Rutgers University Press.

Mercier, H. (2011). What good is moral reasoning? *Mind & Society*, *10*, 131-148.

Michotte, A. (1963). *The perception of causality*. Retrieved from http://psycnet.apa.org/psycinfo/1964-05029-000

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

Mikhail, J. (2009). Is the prohibition of homicide universal? Evidence from comparative criminal law. *Brooklyn Law Review*, *75*, 497-515.

Mill, J. S. (1861). *Utilitarianism*. New York, NY: Oxford University Press.

Miller, R., & Cushman, F. (in press). Moral values and motivations: How special are they? In K. Gray & J. Graham (Eds.), *The atlas of moral psychology: Mapping good and evil*. New York, NY: Guilford Press..

Miller-Young, M (2013, June 10). Pornography can be empowering to women on screen. *The New York Times*. Retrieved from http://www.nytimes.com/roomfordebate/2012/11/11/does-pornography-deserve-its-bad-rap/pornography-can-be-empowering-to-women-on-screen

Mnookin, S. (2012). *The panic virus: The true story behind the vaccine-autism controversy*. New York, NY: Simon & Schuster.

Monroe, A. E., Guglielmo, S., & Malle, B. F. (2012). Morality goes beyond mind perception. *Psychological Inquiry*, *23*, 179-184. doi:10.1080/1047840X.2012.668271

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

Murray, D., & Lombrozo, T. (2016). Effects of manipulation on attributions of causation, free will, and moral responsibility. *Cognitive Science*. Advance online publication. doi:10.1111/cogs.12338

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

Nichols, S., & Folds-Bennett, T. (2003). Are children moral objectivists? Children's judgments about moral and response-dependent properties. *Cognition*, *90*(2), B23-B32.

Niemi, L., & Young, L. (2014). Blaming the victim in the case of rape. *Psychological Inquiry*, *25*, 230-233. doi:10.1080/1047840X.2014.901127

Niemi, L., & Young, L. (2016). When and why we see victims as responsible: The impact of ideology on attitudes toward victims. *Personality and Social Psychology Bulletin*, *42*, 1227-1242. doi:10.1177/0146167216653933

Nisbett, R. E., & Cohen, D. (1996). *Culture of honor: The psychology of violence in the south*. Boulder, CO: Westview Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

Nucci, L., & Turiel, E. (1993). God's word, religious rules, and their relation to Christian and Jewish children's concepts of morality. *Child Development*, *64*, 1475-1491. doi:10.1111/j.1467-8624.1993.tb02965.x

Nucci, L., 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

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.

Parsons, T. (1951). *The social system*. New York, NY: Routledge.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in Cognitive Science*, *2*, 511-527. doi:10.1111/j.1756-8765.2010.01096.x

Peters, J. W. (2016, July 12). Emerging republican platform goes far to the right. *The New York Times*. Retrieved from http://www.nytimes.com/2016/07/13/us/politics/republican-convention-issues.html

Philipps, D. (2016, March 23). North Carolina Bans Local Anti-Discrimination Policies. *The New York Times*. Retrieved from http://www.nytimes.com/2016/03/24/us/north-carolina-to-limit-bathroom-use-by-birth-gender.html

Phillips, J., & Shaw, A. (2015). Manipulating morality: Third-party intentions alter moral judgments by changing causal reasoning. *Cognitive Science*, *39*, 1320-1347. doi:10.1111/cogs.12194

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

Piazza, J., & Sousa, P. (2016). When injustice is at stake, moral judgements are not parochial. *Proceedings. Biological Sciences B*, *282*, 20152037.

Piazza, J., Sousa, P., & Holbrook, C. (2013). Authority dependence and judgments of utilitarian harm. *Cognition*, *128*, 261-270.

Pierce, J. (2001, May 4). Opponents of pornography contend it's not "harmless" or victimless. *Baptist Press*. Retrieved from http://www.bpnews.net/10832/opponents-of-pornography-contend-its-not-harmless-or-victimless

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

Pizarro, D., & 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., Tannenbaum, D., & Uhlmann, E. (2012). Mindless, harmless, and blameworthy. *Psychological Inquiry*, *23*, 185-188. doi:10.1080/1047840X.2012.670100

Preston, S. D., & de Waal, F. B. M. (2001). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1-20. doi:10.1017/S0140525X02000018

Pronin, E., Wegner, D. M., McCarthy, K., & Rodriguez, S. (2006). Everyday magical powers: The role of apparent mental causation in the overestimation of personal influence. *Journal of Personality and Social Psychology*, *91*, 218-231. doi:10.1037/0022-3514.91.2.218

Rachels, J. (1986). *The elements of moral philosophy*. New York, NY: Random House.

Rai, T. S., & Diermeier, D. (2015). Corporations are Cyborgs: Organizations that can think but cannot feel elicit anger as villains but fail to elicit sympathy as victims. *Organizational Behavior and Human Decision Processes*, *126*, 18-26.

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

Rai, T. S., & Fiske, A. P. (2012). Beyond harm, intention, and dyads: Relationship regulation, virtuous violence, and metarelational morality. *Psychological Inquiry*, *23*, 189-193. doi:10.1080/1047840X.2012.670782

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles:(Re) introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review*, *1*, 26-53.

Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The trans-situational influence of social norms. *Journal of Personality and Social Psychology*, *64*, 104-112.

Rest, J. R. (1975). Longitudinal study of the Defining Issues Test of moral judgment: A strategy for analyzing developmental change. *Developmental Psychology*, *11*, 738-748.

Robbins, J., & Rumsey, A. (2008). Introduction: Cultural and linguistic anthropology and the opacity of other minds. *Anthropological Quarterly*, *81*, 407-420. doi:10.1353/anq.0.0005

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

Roberts, C. F., Golding, S. L., & Fincham, F. D. (1987). Implicit theories of criminal responsibility: Decision making and the insanity defense. *Law and Human Behavior*, *11*, 207-232.

Robinson, P. H., & Darley, J. M. (1995). *New Directions in Social Psychology: Justice, liability, and blame: Community views and the criminal law*. Boulder, CO: Westview Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum.

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition*, *108*, 771-780. doi:10.1016/j.cognition.2008.07.001

Rottman, J., Young, L., & Kelemen, D. (2017). The impact of testimony on children's moralization of novel actions. *Emotion*. Advance online publication. http://dx.doi.org/10.1037/emo0000276

Royzman, E., Atanasov, P., Landy, J. F., Parks, A., & Gepty, A. (2014). CAD or MAD? Anger (not disgust) as the predominant response to pathogen-free violations of the divinity code. *Emotion*, *14*, 892-907. doi:10.1037/a0036829

Royzman, E., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, *10*, 296-313.

Rozin, P. (1999). The process of moralization. *Psychological Science*, *10*, 218-221.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*, 574-586. doi:10.1037/0022-3514.76.4.574

Rozin, P., Markwith, M., & Stoess, C. (1997). Moralization and becoming a vegetarian: The transformation of preferences into values and the recruitment of disgust. *Psychological Science*, *8*(2), 67-73.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1 and 2). Cambridge, MA: MIT Press.

Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: What it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological Bulletin*, *139*, 328-351. doi:10.1037/a0029319

Ryan, T. J. (2014). Reconsidering moral issues in politics. *The Journal of Politics*, *76*, 380-397.

Santee, R. T., & Jackson, J. (1977). Cultural values as a source of normative sanctions. *Sociological Perspectives*, *20*, 439-454. doi:10.2307/1388917

Schein, C., Goranson, A., & Gray, K. (2015). The uncensored truth about morality. *The Psychologist*, *28*, 982-985.

Schein, C., & Gray, K. (in press). Moralization: What makes an act wrong? In K. Gray & J. Graham (Eds.), *The atlas of moral psychology: Mapping good and evil*. Guilford Press.

Schein, C., & Gray, K. (2014). The prototype model of blame: Freeing moral cognition from linearity and little boxes. *Psychological Inquiry*, *25*, 236-240. doi:10.1080/1047840X.2014.901903

Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, *41*, 1147-1163. doi:10.1177/0146167215591501

Schein, C., & Gray, K. (2016). Moralization and harmification: The dyadic loop explains how the innocuous becomes harmful and wrong. *Psychological Inquiry*, *27*, 62-65.

Schein, C., Hester, N., & Gray, K. (2016). The visual guide to morality: Vision as an integrative analogy for moral experience, variability and mechanism. *Social and Personality Psychology Compass*, *10*, 231-251. doi:10.1111/spc3.12247

Schein, C., Ritter, R., & Gray, K. (2016). Harm mediates the disgust-immorality link. *Emotion*, *16*, 862-876.

Schelling, T. C. (1981). *The strategy of conflict* (Reprint ed.). Cambridge, MA: Harvard University Press.

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

Schwartz, S. H. (1999). A theory of cultural values and some implications for work. *Applied Psychology*, *48*, 23-47.

Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, *53*, 550-562. doi:10.1037/0022-3514.53.3.550

Scott, S., Inbar, Y., & Rozin, P. (2016). Evidence for absolute moral opposition to genetically modified food in the United States. *Perspectives on Psychological Science*, *11*, 315-324.

Seligman, M. E. (1971). Phobias and preparedness. *Behavior Therapy*, *2*, 307-320.

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

Sheskin, M., & Baumard, N. (2016). Switching away from utilitarianism: The limited role of utility calculations in moral judgment. *PLoS ONE*, *11*(8), e0160084. doi:10.1371/journal.pone.0160084

Sheskin, M., & Santos, L. (2012). The evolution of morality: Which aspects of human moral concerns are shared with nonhuman primates. In J. Vonk & T. K. Shackelford (Eds.), *The Oxford handbook of comparative evolutionary psychology* (pp. 434-450). New York, NY: Oxford University Press.

Shultz, T. R., Schleifer, M., & Altman, I. (1981). Judgments of causation, responsibility, and punishment in cases of harm-doing. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *13*, 238-253. doi:10.1037/h0081183

Shweder, R. A. (2012). Relativism and universalism. In D. Fassin (Ed.), *A companion to moral anthropology* (pp. 85-102). Hoboken, NJ: John Wiley.

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 A. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119-169). New York, NY: Routledge.

Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental Psychology*, 5(1), 136-150.

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

Singer, P. (1981). *The expanding circle*. Oxford, UK: Clarendon Press.

Sinnott-Armstrong, W. (2016). The disunity of morality. In S. M. Liao (Ed.), *Moral brains: The neuroscience of morality* (pp. 331-354). Oxford, UK: Oxford University Press.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, 88, 895-917. doi:10.1037/0022-3514.88.6.895

Skitka, L. J., & Houston, D. A. (2001). When due process is of no consequence: Moral mandates and presumed defendant guilt or innocence. *Social Justice Research*, 14, 305-326.

Skitka, L. J., Washburn, A. N., & Carsel, T. S. (2015). The psychological foundations and consequences of moral conviction. *Current Opinion in Psychology*, 6, 41-44.

Skitka, L. J., & Wisneski, D. C. (2011). Moral conviction and emotion. *Emotion Review*, 3, 328-330.

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, 52, 1333-1336.

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology*, 21(1), 18-29.

Smetana, J. G., & Killen, M. (2008). Moral cognition, emotions, and neuroscience: An integrative developmental view. *International Journal of Developmental Science*, 2, 324-339.

Smith, A. (1812). *The theory of moral sentiments*. Retrieved from http://books.google.com/books?hl=en&lr=&id=d-UUAAAA QAAJ&oi=fnd&pg=PP18&dq=adam+smith+theory+moral+s entiment&ots=mjgFFCLLgg&sig=OCnnF0s8QzbIGhdXaogP haboXh0

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Smith, E. E., & Medin, D. L. (2002). The exemplar view. In D. J. Levitin (Ed.), *Foundations of cognitive psychology: Core readings* (pp. 277-292). Cambridge, MA: The MIT Press.

Spellman, B. A., & Gilbert, E. A. (2014). Blame, cause, and counterfactuals: The inextricable link. *Psychological Inquiry*, 25, 245-250. doi:10.1080/1047840X.2014.901708

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 39-67). New York, NY: Cambridge University Press.

Spivey, M. J., & Dale, R. (2004). The continuity of mind: Toward a dynamical account of cognition. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 45, pp. 87-142). San Diego, CA: Elsevier.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27, 76-105.

Sripada, C. S., & Stich, S. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Culture and cognition* (Vol. 2, pp. 280-301). Oxford, UK: Oxford University Press.

Stahl, T., Zaal, M. P., & Skitka, L. J. (2016). Moralized rationality: Relying on logic and evidence in the formation and evaluation of belief can be seen as a moral issue. *PLoS ONE*, 11(11), e0166332.

Staub, E. (1992). *The roots of evil: The origins of genocide and other group violence* (Rev. ed.). Cambridge, UK: Cambridge University Press.

Steinberg, L., & Scott, E. S. (2003). Less guilty by reason of adolescence: Developmental immaturity, diminished responsibility, and the juvenile death penalty. *American Psychologist*, 58, 1009-1018.

Stenner, K. (2005). *The authoritarian dynamic*. Cambridge, UK: Cambridge University Press.

Stout, H. (2013, October 21). Not just a hot cup anymore. *The New York Times*. Retrieved from http://www.nytimes.com/2013/10/21/booming/not-just-a-hot-cup-anymore.html

Strickland, B., Fisher, M., & Knobe, J. (2012). Moral structure falls out of general event structure. *Psychological Inquiry*, 23, 198-205. doi:10.1080/1047840X.2012.668272

Suarez, E., & Gadalla, T. M. (2010). Stop blaming the victim: A meta-analysis on rape myths. *Journal of Interpersonal Violence*, 25, 2010-2035. doi:10.1177/0886260509354503

Sunstein, C. R., Kahneman, D., & Schkade, D. (1998). Assessing punitive damages (with notes on cognition and valuation in law). *The Yale Law Journal*, 107, 2071-2153. doi:10.2307/797417

Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition*, 119, 454-458. doi:10.1016/j.cognition.2011.01.018

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. (2011). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology*, 47, 1249-1254.

Taylor, J. R. (2003). *Linguistic categorization*. Oxford, UK: Oxford University Press.

Tetlock, P. E., Kristel, O. V., Beth, S., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, 78, 853-870. doi:10.1037/0022-3514.78.5.853

Thagard, P. (2002). *Coherence in thought and action*. Cambridge, MA: The MIT press.

Thelen, E. (1996). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: The MIT Press.

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

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 155-243). Chicago, IL: University of Chicago Press.

Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and structure. *Psychological Review*, 120, 65-84. doi:10.1037/a0030778

Uhlmann, E. L., & Zhu, L. (2013). Acts, persons, and intuitions person-centered cues and gut reactions to harmless

transgressions. *Social Psychological and Personality Science*, *5*, 279-285. doi:10.1177/1948550613497238

Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, *17*, 476-477. doi:10.1111/j.1467-9280.2006.01731.x

Vallacher, R. R., Read, S. J., & Nowak, A. (2002). The dynamical perspective in personality and social psychology. *Personality and Social Psychology Review*, *6*, 264-273. doi:10.1207/S15327957PSPR0604_01

Van Bavel, J. J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: from dual processes to dynamic systems. *Current Opinion in Psychology*, *6*, 167-172.

Van de Vondervoort, J. W., & Hamlin, J. K. (2016). Evidence for intuitive morality: Preverbal infants make sociomoral evaluations. *Child Development Perspectives*, *10*, 143-148.

van Leeuwen, F., & Park, J. H. (2009). Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences*, *47*, 169-173. doi:10.1016/j.paid.2009.02.017

van Leeuwen, F., Park, J. H., Koenig, B. L., & Graham, J. (2012). Regional variation in pathogen prevalence predicts endorsement of group-focused moral concerns. *Evolution and Human Behavior*, *33*, 429-437. doi:10.1016/j.evolhumbehav.2011.12.005

van Zomeren, M., Postmes, T., & Spears, R. (2012). On conviction's collective consequences: Integrating moral conviction with the social identity model of collective action. *The British Journal of Social Psychology*, *51*, 52-71. doi:10.1111/j.2044-8309.2010.02000.x

Wainryb, C., Shaw, L. A., Langley, M., Cottam, K., & Lewis, R. (2004). Children's thinking about diversity of belief in the early school years: Judgments of relativism, tolerance, and disagreeing persons. *Child Development*, *75*, 687-703.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology*, *5*(7), e184.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*, 1301-1303. doi:10.1126/science.1121448

Weeks, L. (2012, August 24). Lance Armstrong: When a hero lets us down. *npr*. Retrieved from http://www.npr.org/2012/08/24/160006508/lance-armstrong-when-a-hero-lets-us-down

Wegner, D. M., & Gray, K. (2016). *The mind club: Who thinks, what feels, and why it matters*. New York, NY: Viking.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*, 480-492. doi:10.1037/0003-066X.54.7.480

Weiner, B. (1993). On sin versus sickness: A theory of perceived responsibility and social motivation. *American Psychologist*, *48*, 957-965.

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

Wisneski, D. C., Lytle, B. L., & Skitka, L. J. (2009). Gut reactions: Moral conviction, religiosity, and trust in authority. *Psychological Science*, *20*, 1059-1063. doi:10.1111/j.1467-9280.2009.02406.x

Wisneski, D. C., & Skitka, L. (2017). Moralization through moral shock: Exploring emotional antecedents to moral conviction. *Personality and Social Psychological Bulletin*, *43*(2), 139-150. doi:10.1177/0146167216676479

Wittgenstein, L. (1953). *Philosophical investigations: The German text, with a revised English translation* (G. E. M. Anscombe, Trans.). Oxford, UK: Blackwell.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34.

Woodward, A. L. (2005). The infant origins of intentional understanding. *Advances in Child Development and Behavior*, *33*, 229-262.

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*, 8235-8240.

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, *7*(1), 1-10. doi:10.1080/17470919.2011.569146

Young, L., & Saxe, R. (2010). It's not just what you do, but what's on your mind: A review of Kwame Anthony Appiah's "Experiments in Ethics." *Neuroethics*, *3*, 201-207.

Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, *120*, 202-214. doi:10.1016/j.cognition.2011.04.005

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151-175.