

The Myth of Harmless Wrongs in Moral Cognition: Automatic Dyadic Completion From Sin to Suffering

Kurt Gray and Chelsea Schein
University of North Carolina, Chapel Hill

Adrian F. Ward
University of Colorado, Boulder

When something is wrong, someone is harmed. This hypothesis derives from the theory of dyadic morality, which suggests a moral cognitive template of wrongdoing agent and suffering patient (i.e., victim). This dyadic template means that victimless wrongs (e.g., masturbation) are psychologically incomplete, compelling the mind to perceive victims even when they are objectively absent. Five studies reveal that dyadic completion occurs automatically and implicitly: Ostensibly harmless wrongs are perceived to have victims (Study 1), activate concepts of harm (Studies 2 and 3), and increase perceptions of suffering (Studies 4 and 5). These results suggest that perceiving harm in immorality is intuitive and does not require effortful rationalization. This interpretation argues against both standard interpretations of moral dumbfounding and domain-specific theories of morality that assume the psychological existence of harmless wrongs. Dyadic completion also suggests that moral dilemmas in which wrongness (deontology) and harm (utilitarianism) conflict are unrepresentative of typical moral cognition.

Keywords: morality, politics, moral foundations, values, moral dyad

“In my opinion neither the plague, nor war, nor small-pox, nor similar diseases, have produced results so disastrous to humanity as the pernicious habit of Onanism; it is the destroying element of civilized societies.” (Dr. Adam Clarke on masturbation, quoted in Kellogg, 1890, p. 233)

Irreparable harm from masturbation may seem far-fetched, but this morally contentious act has long been linked to suffering through abnormal development, unnatural baldness, and even blindness and paralysis (Kellogg, 1890, pp. 249–254; Qur’an 23:5–23:7). Many other victimless acts (e.g., homosexuality, dietary choices, private drug use) have also been thought to harm the self (Kellogg, 1890), specific others (Bryant, 1977), and society at large (*Hollingsworth v. Perry*, 2013). For example, homosexuality is perceived by some to harm children via physical harm, mental suffering, or spiritual corruption (Bryant, 1977). Whether these acts cause objective harm is debated, but it appears that judgments of immorality are tied to perceived harm. Such perceived harm has often been interpreted as a product of effortful, conscious rationalization (Haidt, 2001), but we suggest an alternative possibility: Immorality automatically activates perceptions of harm, consistent with a dyadic moral template that binds together immoral agents with suffering victims (Gray, Waytz, & Young, 2012). We conducted five studies to test the implicit persistence of harmed victims in immoral but objectively harmless acts.

Moral Dumbfounding and Perceived Harm

Western philosophy and classic psychology have long discussed the link between moral judgments and harm (Kohlberg, 1969; Mill, 1863/2008; Piaget, 1932), but the phenomenon of moral dumbfounding seems to suggest that this link is easily broken. In an oft-cited but unpublished study (Haidt, Bjorklund, & Murphy, 2000), participants judged as immoral some scenarios engineered to be objectively harmless (e.g., consensual incest). These judgments of wrongness remained even after researchers explicitly disqualified harm-based explanations (e.g., disfigured children) with scenario facts (e.g., birth control is used). This phenomenon is called moral dumbfounding, because participants are rendered “dumb” to explain their enduring moral judgments without referencing the experimentally disallowed concept of harm.

The dissociation of objective harm from immorality (i.e., the existence of harmless wrongs) has inspired theorizing that some moral concerns are independent from harm. In particular, accounts of moral pluralism suggest that violations related to divinity/purity (e.g., eating a dead dog) and community/loyalty (e.g., flag burning) are processed by encapsulated moral mechanisms cognitively unconnected to harm (Graham et al., 2013; Haidt & Joseph, 2007; Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011). For example, moral foundations theory (Graham et al., 2013) assumes the harmlessness of purity violations by including questionnaire items such as “people should not do things that are disgusting, even if no one is harmed” (Graham, Haidt, & Nosek, 2009).

It is clear that purity violations (e.g., deviant sexuality) are descriptively different from canonical harmful violations (e.g., murder) and are relatively less likely to cause direct physical harm. However, some researchers have taken an additional inferential step by assuming that scenarios “carefully written to be harmless” (Haidt et al., 2000, p. 6) are actually seen to be harmless by participants. If this assumption is true, it would imply that reports

This article was published Online First March 17, 2014.

Kurt Gray and Chelsea Schein, Department of Psychology, University of North Carolina, Chapel Hill; Adrian F. Ward, Department of Marketing, Leeds School of Business, University of Colorado, Boulder.

Kurt Gray and Chelsea Schein contributed equally to this work.

Correspondence concerning this article should be addressed to Kurt Gray, Department of Psychology, University of North Carolina, Chapel Hill, Chapel Hill, NC 27599. E-mail: kurtgray@unc.edu

of harm in impurity stem from effortful rationalization of moral intuitions (Haidt, 2001). In other words, truly “harmless immorality” would suggest that although moral judgments are automatic and intuitive, perceptions of harm are not. However, it is unclear whether participants reading these scenarios share the perceptions of harmlessness of the researchers who write them: Participants offer harm-based explanations for scenario wrongness, even when researchers deny its presence, and are rendered dumb only when they are forbidden from mentioning harm (Haidt et al., 2000).

Moral dumbfounding may demonstrate that moral judgment can be independent from “objective” harm (but see Jacobson, 2012), but moral judgment may nevertheless be linked to *perceptions of subjective harm*. Studies across psychology have long documented the separation between subjective experience and objective fact. People perceive lines differing in length even though lines are objectively identical (Müller-Lyer, 1889), perceive plane travel as dangerous although it is safer than car travel (Ropeik, 2010), and can even perceive the presence of a limb after its amputation (Shukla, Sahu, Tripathi, & Gupta, 1982). Indeed, one of the key tenets of psychology is that perception is dissociable from objective reality (James, 1890). In moral dumbfounding, people may still intuitively perceive harm despite the objective disavowals of experimenters. Just as safety statistics cannot prevent you from feeling uneasy when your plane twists in turbulence, so too might the disavowals of experimenters be unable to prevent people from intuitively perceiving harm. Imagine that a tarantula—guaranteed harmless—is placed on your face; you would likely sweat and twitch and try to escape, despite any objective assurances of its harmlessness (Gendler, 2008). Across diverse nonmoral areas, intuitions of subjective harm persist despite objective harmlessness; we suggest the same is possible in moral cognition.

Research on mind perception underscores the subjectivity of harm. Harm depends upon perceiving a victimized mind, and other minds are ultimately unknowable and therefore ambiguous (Chalmers, 1997; Epley & Waytz, 2009; Haslam, Loughnan, Kashima, & Bain, 2008; Waytz, Cacioppo, & Epley, 2010). People can fail to see harm in cases of genocide (Castano & Giner-Sorolla, 2006; Jahoda, 1998), child slavery (Gorney, 2011), or torture (Gray & Wegner, 2010b; Greenberg & Dratel, 2005) simply by stripping others of mind (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Haslam, 2006). Conversely, people can see harm in apparently victimless acts by ascribing more mind to animals (Bastian, Costello, Loughnan, & Hodson, 2012), fetuses (Gray, Gray, & Wegner, 2007), nature (Tam, Lee, & Chao, 2013), vegetative patients (Gray, Knickman, & Wegner, 2011), and robots (Gray & Wegner, 2012; Ward, Olsen, & Wegner, 2013). Motivation can also alter perceived harm, as people strip minds from those they hate (Castano & Giner-Sorolla, 2006; Goff, Eberhardt, Williams, & Jackson, 2008; Haslam, 2006; Osofsky, Bandura, & Zimbardo, 2005) and confer them on those they love (Gardner & Knowles, 2008; Waytz, Epley, & Cacioppo, 2010; Waytz, Gray, Epley, & Wegner, 2010). In judgments and behavior, the perception of harm (i.e., seeing suffering minds) often matters more than objective reality. In fact, mind perception is sufficiently labile that it may make more sense to discuss “ostensibly harmless” scenarios rather than “objectively harmless” scenarios, because harm can never be assessed completely objectively.

Beyond the general ambiguity of mind perception, the subjective nature of harm within morality is bolstered by the subjective

nature of morality. Although philosophers debate the existence of objective moral truths (Appiah, 2008), psychologists typically recognize that morality is a matter of perception (Haidt, 2012). Moral judgments are made quickly and effortlessly (Cushman, Young, & Hauser, 2006; Gigerenzer, 2008; Haidt, 2001) and are often insensitive to objective, rational considerations for moral violations on the basis of feelings of retribution (Carlsmith, 2006; Greene & Cohen, 2004) and perceptions of mind (Gray & Wegner, 2011b) rather than of more rational concerns, such as deterrence. In sum, past research has demonstrated that judgments of both morality and harm are matters of perception, and we simply suggest that these two perceptions are intertwined such that moral judgments activate intuitions about harm.¹

Moral Dyad and Dyadic Completion

Psychology has long suggested general associations between immorality and harm (Piaget, 1932; Preston & de Waal, 2001; Smith, 1759/1882; Sousa, Holbrook, & Piazza, 2009; Turiel, 1983), but a new framework suggests a specific and persistent cognitive link between moral wrongs and perceived harm. Dyadic morality, a concept grounded in the cognitive psychology of concepts, suggests that morality is understood through a harm-based template of two perceived minds: a wrongdoing agent (A) acting upon a suffering patient (P); [A–P] (Gray, Waytz, & Young, 2012; Gray, Young, & Waytz, 2012). This dyadic template combines two dimensions of mind perception—agency and experience (Gray et al., 2007; Gray, Jenkins, Heberlein, & Wegner, 2011)—into a causal structure that grows out of the frequency, universality, and affective power of harm (Davis, 1996; Decety & Cacioppo, 2012; Decety & Meyer, 2008; Haidt, Koller, & Dias, 1993), as well as the dyadic nature of language, action, and thought (Brown & Fish, 1983; Strickland, Fisher, & Knobe, 2012). The three components of the dyad—intentional moral agent, causation, and suffering moral patient—are three broad elements highlighted by moral psychology (Hauser, Young, & Cushman, 2007; Mikhail, 2007), psychodynamic theory (Karpman, 1968), the law (Hart & Honoré, 1985), and everyday folk psychology (Guglielmo, Monroe, & Malle, 2009).

This dyadic harm-based template of wrongdoer + victim is not a static representation but instead serves as a cognitive working model with both bottom-up and top-down effects (Craik, 1967). These bottom-up effects are relatively well known, as many studies have documented how the addition of agency, causation, and suffering can increase the severity of moral judgment (Cushman, 2008; Gray & Schein, 2012; Spranca, Minsk, & Baron, 1991; Weiner, 1995). Top-down effects suggest that the dyadic template should shape the perception of morally ambiguous scenarios, just as stereotypes shape the perception of racially ambiguous targets (Peery & Bodenhausen, 2008). That is, when moral judgments are triggered—whether through norm violations (Monroe, Guglielmo, & Malle, 2012; Nichols, 2002), disgusting smells (Schnall, Haidt, Clore, & Jordan, 2008), or affronts to God or country (Graham & Haidt, 2010; Graham et al., 2009)—this dyadic template should compel the perception of agents, causation, or patients, even when

¹ Of course, even subjective perceptions are likely to feel objective to perceivers themselves (Goodwin & Darley, 2008).

these factors are ostensibly absent. This process is called *dyadic completion*, and it has three flavors, one for each missing element of the dyad.

Agentic dyadic completion occurs for isolated moral patients [–P], compelling the perception of intentional agents to blame for suffering, including God (Gray & Wegner, 2010a), animals (Oldridge, 2004), and other people (Knobe, 2003). It can occur even in the face of potential harm, where the prospect of natural disasters make profiting third parties seem evil (Inbar, Pizarro, & Cushman, 2012; Tannenbaum, Uhlmann, & Diermeier, 2011).

Causal dyadic completion occurs for disjoint agents and patients [A P], compelling a causal link between them. This is best demonstrated by studies of culpable causation in which a drug dealer rushing home to hide cocaine is ascribed more blame for a car crash than someone rushing home for a more innocuous reason is (Alicke, 1992, 2000). Causal dyadic completion is so powerful that people will even believe in voodoo to causally connect their cruel intentions with another person's suffering (Pronin, Wegner, McCarthy, & Rodriguez, 2006).

Most important for the current paper is *patientic dyadic completion*, which should occur for isolated moral agents [A—], compelling the perception of suffering moral patients resulting from immoral deeds. In other words, moral wrongs—even ostensibly victimless ones—should lead to the perception of harm. As a rough analogy, consider the phenomenon of visual completion in the Kanizsa triangle, in which the presence of one triangle (an intentional agent) and the surrounding shapes (a moral context) compel the perception of a second complementary triangle (a suffering patient). See Figure 1. Of course, visual completion is a much lower level process, but we suggest that the firsthand phenomenological experience is similar (Gray & Wegner, 2013). Just as people cannot help but see the second triangle despite its objective absence, we suggest, people cannot help but see the presence of harm in harmless wrongs.

Research reveals an explicit link between immorality and perceived harm (e.g., Ditto & Liu, 2011) consistent with patientic dyadic completion. People see victims or suffering in proportion to the severity or intentionality of various moral transgressions

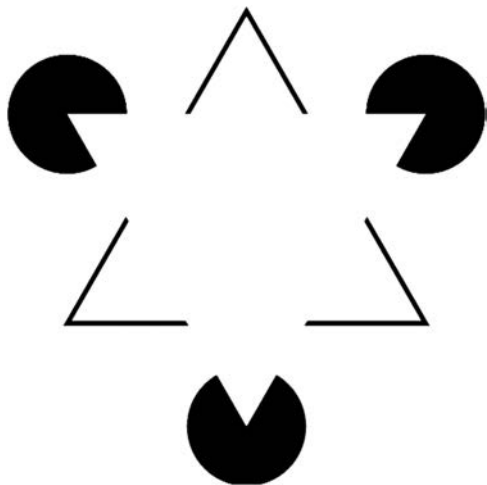


Figure 1. The Kanizsa triangle: an example of automatic visual completion.

(DeScioli, 2008; DeScioli, Gilbert, & Kurzban, 2012; Gutierrez & Giner-Sorolla, 2007; Haidt & Hersh, 2001; Ward et al., 2013). For example, intentional harms are seen to cause more harm (Gray & Wegner, 2008) and to be overall worse than unintentional harms (Ames & Fiske, 2013). DeScioli et al. (2012) cleverly termed this phenomenon the *indelible victim effect*, but we use the term dyadic completion because it refers to a broader psychological process. Such completion can even translate into experience, as shocks administered by malicious agents actually hurt more than those administered accidentally or by a computer (Gray, 2012; Gray & Wegner, 2008).² Patientic dyadic completion suggests that harmless wrongs or victimless crimes are psychologically incomplete, and so the mind automatically fills in the missing moral patient.

The Current Research

Past work on dyadic completion has used explicit measures, and so the link between wrong and harm has often been interpreted as effortful rationalization. We instead suggest it occurs automatically and use implicit measures to test whether ostensibly harmless scenarios activate the concept of harm more than other negative nonmorally relevant concepts do (Studies 2 and 3). As the Merriam-Webster's dictionary definition of *harm* (n.d.) includes "something that causes someone or something to be hurt," we also assess whether ostensibly harmless scenarios involve perceived victims (i.e., "someone or something"; Study 1) and perceptions of physical and emotional suffering ("hurt"; Studies 4 and 5). We predict that judgments of immorality will be implicitly associated with perceptions of harm even in ostensibly victimless moral transgressions.

Study 1: Time Pressure and Perceived Victimhood

It is clear that people can use effortful reasoning to justify their moral judgments, but dyadic completion also suggests that people link wrongness to harm implicitly, without requiring mental resources needed for effortful justification. In this study, participants rated the wrongness and perceived victimhood of ostensibly harmless scenarios across high and low time pressure. If perceived harm is a product of effortful rationalization, victimhood ratings should decrease under the mental constraints of time pressure. However, if perceived harm is intuitively associated with immorality, victimhood ratings should be similar or even enhanced under the time pressure. Consistent with the latter possibility, past work has found that ratings of harm are relatively unaffected by cognitive load (Gutierrez & Giner-Sorolla, 2007; Wright & Baril, 2011). However, this study differs from past work by asking specifically about the presence of victims (i.e., moral patients), which a dyadic template suggests should be especially salient in moral wrongs.

Method

Recruitment and exclusion criterion. In this and all subsequent studies (except Study 2), participants were recruited through

² Dyadic completion may also function in the positive domain, as good intentions improve the experience of massages and food (Gray, 2012).

MTurk. Across all studies, participants earned \$0.15 to \$1.00, depending on the length of the study. Internet samples are frequently used in psychological research (Skitka & Sargis, 2006), and MTurk recruitment maintains reliability equal to lab-based populations while providing greater diversity (Buhrmester, Kwang, & Gosling, 2011). However, MTurk samples are more likely than lab-based samples to encounter problems due to technological glitches and/or inattentive participants (Goodman, Cryder, & Cheema, 2013). Participants were excluded because they failed attention checks (Kapelner & Chandler, 2010; Oppenheimer, Meyvis, & Davidenko, 2009) or because they failed to follow instructions.

Participants. One hundred and three participants completed this study through MTurk. Ten participants failed the instructional manipulation check, and another 11 participants reported that neutral scenarios such as eating toast or folding a letter were immoral, leaving 82 participants (37% female, $M_{\text{age}} = 39$ years, 52% liberal, all from the United States).

Procedure. Each participant rated the morality of 12 different actions, including four ostensibly victimless but impure moral violations (masturbating to a picture of one's dead sister, watching animals have sex to become sexually aroused, having sex with a corpse, covering a Bible with feces), four harmful actions (sticking a stranger with a pin, insulting an overweight colleague, kicking a dog hard, beating one's wife) and four neutral scenarios (eating toast, riding the bus, folding a letter, reading an article; see the Appendix for all scenarios). These scenarios were adapted from scenarios used to validate accounts of harm-independent immorality (Graham et al., 2009; Haidt et al., 1993). After reading each scenario, participants rated, on 5-point scales, the action's moral wrongness from *Not Wrong at All* (1) to *Extremely Wrong* (5) and whether the action has a victim (or victims) from *Definitely Not* (1) to *Definitely Yes* (5). Wrongness and victim ratings were collapsed by scenario type: impure ($\alpha > .70$), harmful ($\alpha > .71$), and neutral ($\alpha = .45$).³

To manipulate participants' ability to engage in effortful justification, we randomly assigned participants to either the *time pressure* or the *ample time* condition. Time pressure is a well-validated manipulation used in many studies of cognitive load (Svenson & Maule, 1993). Time pressure impairs people's capacity to correct judgments (Gilbert & Gill, 2000), increases egocentric bias (Epley, Keysar, Van Boven, & Gilovich, 2004), increases reliance on ethnic stereotypes (Kruglanski & Freund, 1983), and decreases accuracy for most decision strategies (Payne & Bettman, 2004).

Participants in the time pressure condition were instructed to respond with their gut reaction, going as quickly as possible, and were given only 7 seconds to read each scenario and provide their answers. A countdown clock was displayed prominently on the bottom of the screen to increase pressure and divert cognitive resources. Participants in the ample time condition not only were given unlimited time but were able to answer questions only after a 7-s delay (similar to Paxton, Ungar, & Greene, 2012), during which they were instructed to think carefully.

As an independent assessment of whether the capacity for effortful reasoning differed across conditions, participant then completed 12 arithmetic questions (e.g., $137 + 53$) by selecting one of two answers (e.g., 200 or 190). As expected, participants in the time pressure condition answered fewer arithmetic questions

correctly ($M = 8.83$, $SD = 2.05$) than those in the time delay condition did ($M = 11.37$, $SD = 0.95$), $t(80) = 7.44$, $p < .001$.

Finally, participants reported demographic information (gender, age, political orientation, and country). To test whether our purity scenarios were consistent with past research (e.g., Graham et al., 2009), we examined correlations between their perceived wrongness and political affiliation. As predicted, politics correlated with both the perceived immorality, $r(80) = .25$, $p = .03$, and victimhood, $r(80) = .23$, $p = .04$, of purity violations but not with ratings of harm scenarios or victims in neutral scenarios ($r_s < |.14|$, $p_s > .23$).

Results and Discussion

To examine the role of effortful reasoning in perceived victimhood, we conducted a 3 (scenario: impure, harm, neutral) \times 2 (time condition: time pressure, ample time) within/between analysis of variance (ANOVA). It revealed an expected significant main effect of scenario, such that harmful scenarios ($M = 4.82$, $SD = 0.26$) were seen to have significantly more victims than impure scenarios ($M = 2.46$, $SD = 1.05$; $p < .001$), which had more victims than the neutral scenarios ($M = 1.02$, $SD = 0.13$; $p < .001$), $F(1, 80) = 858.50$, $p < .001$, $\eta^2 = .92$. There was also a significant main effect of time condition, such that more victims were seen under time pressure ($M = 2.91$, $SD = 0.33$) than when participants had ample time ($M = 2.66$, $SD = 0.33$), $F(1, 80) = 11.71$, $p = .001$, $\eta^2 = .13$.

Of importance, these main effects were qualified by a significant interaction, $F(2, 160) = 15.43$, $p < .001$, $\eta^2 = .16$, such that participants saw significantly more victims in impure scenarios under time pressure ($M = 2.93$, $SD = 1.20$) than when they had ample time ($M = 2.09$, $SD = 0.74$; $p < .001$). See Figure 2 for a display of the means. Perceptions of victims in harm scenarios did not vary significantly between time pressure ($M = 4.77$, $SD = 0.31$) and ample time ($M = 4.86$, $SD = 0.21$; $p = .11$), nor did perceptions vary significantly in the neutral scenarios ($p = .65$), likely due to ceiling effects (for harm scenarios) and floor effects (for neutral scenarios). Because significant correlations existed between politics and moral judgment, we examined an additional 3 \times 2 model in which politics was added as a covariate; politics emerged as marginally significant, $F(1, 79) = 3.27$, $p = .08$ —more conservatism, more victim perception—and had a marginally significant interaction with scenario, $F(2, 158) = 2.77$, $p = .07$, but it did not meaningfully affect the significance of other results.

Similar results were obtained when the data were analyzed through a regression with condition, immorality, and politics predicting victim ratings in impure scenarios. Victim ratings were significantly predicted by ratings of immorality, $\beta = .25$, $t(78) = 2.43$, $p = .02$, and time pressure, $\beta = -.35$, $t(78) = -3.55$, $p = .001$, but not politics, $\beta = .12$, $t(78) = 1.21$, $p = .23$. This is an understandable result, given that political affiliation acts directly on ratings of immorality (Graham et al., 2009) and immorality is included in the model.

³ The alpha for neutral victim ratings is low because, for some reason, "folding a letter" correlated poorly with other scenarios. Without this story the Cronbach's alpha was higher ($\alpha = .69$). Analysis reported here exclude this scenario; however, results are nearly identical with it included.

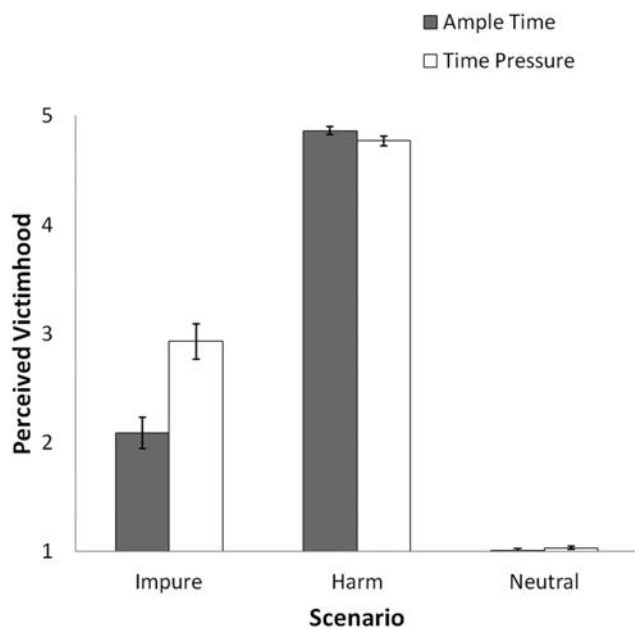


Figure 2. Impact of time pressure on perception of victims by scenario (Study 1). Error bars represent standard error.

Consistent with dyadic morality, victims are subjectively perceived in immoral scenarios designed to be objectively harmless. These victims are perceived even more when effortful reasoning is inhibited through time pressure, suggesting that perceptions of harm in immorality are relatively automatic and intuitive. Of course, 7 seconds is still long enough to allow for some effortful reasoning, and so the next study used an even quicker implicit measure.

Study 2: Misattributions of Harm

The affective misattribution paradigm (AMP) measures implicit affective and semantic associations to diverse stimuli (Gawronski & Ye, 2014; Payne, Cheng, Govorun, & Stewart, 2005). Here, we use it to test whether immoral but ostensibly victimless scenarios implicitly activate the concept of harm. Of importance, we also assess associations of general negativity, which often go untested in studies of moral cognition. To ensure that “harm” is not simply a synonym for general negativity, we controlled for affect congruence by including nonmoral negative scenarios and/or measurements in this and other studies. In this study, we use perceptions of sadness as a negative control, and in later studies we use perceptions of failure (Study 3), attractiveness (Study 4), and boredom (Study 5).

As immorality is affectively negative, we predict that both immoral scenarios (e.g., masturbating to a picture of your dead sister) and negative but nonmoral scenarios (e.g., a child losing a stuffed animal) would activate general negativity. However, we predict that only immoral scenarios will simultaneously activate the concepts of immorality, harm, and negativity. In other words, many things (including immorality) are negative, but only immorality should be negative and wrong and harmful.

Method

Participants. Fifty-six United States college students (63% female, $M_{\text{age}} = 19$ years, 31% liberal) completed this within-subjects study for partial fulfillment of course credit.

Procedure. Participants first read 12 scenarios, including the four purity violations (Defile Corpse, Sister Masturbate, Animal Sex, Bible Feces) and four neutral scenarios (Read Text, Ride Bus, Fold Paper, Eat Toast) from Study 1. They also read four nonmoral negative control actions, including a child losing her stuffed bear (Lose Teddy), a student failing an exam (Fail Exam), a romantic partner leaving (Partner Leave), and a pet gone missing (Cat Missing). Participants were instructed to remember the two-word title of each scenario, as implicit tasks such as the AMP proceed too fast to allow detailed reading. To facilitate a “moral mindset,” participants then categorized the actions as immoral or not immoral as quickly as they could.

Participants then completed a modified AMP (Payne et al., 2005), in which the title of a scenario (e.g., “Bible Feces”) was presented on the screen for 250 ms, followed by a blank screen for 125 ms, and then a Chinese character for 250 ms, followed finally by a static screen mask. Depending on the block, participants (all non-Chinese speakers) were instructed to select whether they thought the meaning of the Chinese character was harmful/sad/wrong or not. Participants were instructed to ignore the scenario title preceding the character, as consistent with typical AMP instructions, and to focus instead on their gut reaction to the Chinese character. Importantly, because the Chinese character is ambiguous in meaning, activated concepts bleed through in these ratings. Semantic activation of concepts in the AMP has been validated across a number of concepts, including animacy (Deutsch & Gawronski, 2009), sexual interest (Imhoff, Schmidt, Bernhardt, Dierksmeier, & Banse, 2011), and personality (Sava et al., 2012). Here, we use it to assess activations of immorality, harm, and more general negativity. Each of the 12 scenarios appeared in each block four times in random order. After completing the AMP, participants completed demographics information.

Results and Discussion

As in Study 1, responses were collapsed by scenario type: impure ($\alpha > .74$), neutral ($\alpha > .76$), and nonmoral negative ($\alpha > .46$).⁴ A 3 (rating: harmful, sad, wrong) \times 3 (scenario: impure, negative control, neutral control) within-subject ANOVA revealed a significant main effect of rating, $F(2, 110) = 11.32, p < .001, \eta^2 = .17$, and scenario, $F(2, 110) = 51.74, p < .001, \eta^2 = .49$. See Figure 3 for a display of the means by condition.

The ANOVA also revealed a significant interaction, $F(4, 220) = 18.96, p < .001, \eta^2 = .26$, such that impure scenarios significantly activated harmful ($M = .70, SD = .24$) more than did either the nonmoral negative ($M = .41, SD = .25$) or the neutral

⁴ The alpha for this type is low because Partner Leave correlated poorly with other scenarios (some may have viewed the wife’s action as abandonment and therefore immoral). Without this story, the Cronbach’s alpha was higher ($\alpha > .65$). Analysis reported here excludes Partner Leave; however, results are nearly identical with this scenario included.

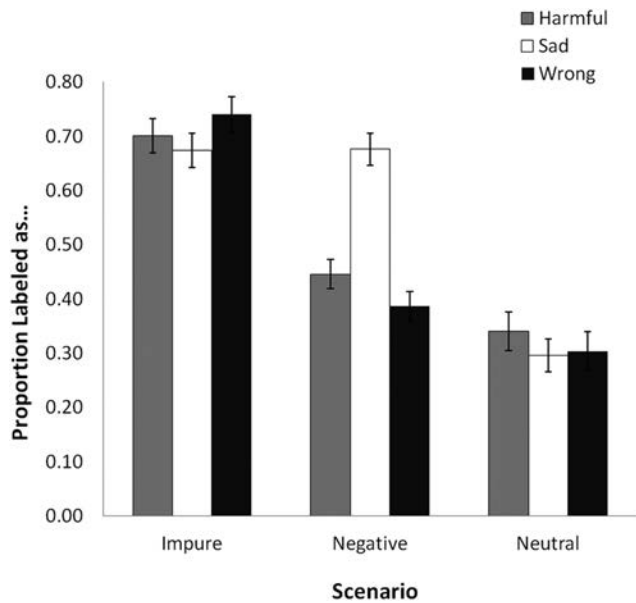


Figure 3. Ratio of Chinese letters labeled as harmful, sad, or wrong based on scenario (Study 2). Error bars represent standard error.

($M = .34$, $SD = .27$) scenarios ($p < .001$).⁵ Consistent with being seen as immoral, impure scenarios also activated wrong ($M = .74$, $SD = .25$) significantly more than did the negative ($M = .32$, $SD = .26$) and neutral scenarios ($M = .30$, $SD = .27$; $p < .001$). As predicted, impure scenarios were not significantly different in sadness ($M = .67$, $SD = .24$) than the negative scenarios ($M = .67$, $SD = .23$; $p = .85$), but both were significantly higher than neutral scenarios ($M = .30$, $SD = .23$; $p < .001$).

These analyses reveal that impure scenarios activated not only the concept of moral wrongness and nonmoral negativity (i.e., sadness) but also harm, just as dyadic completion predicts. Politics did not have a significant between-subjects main effect ($p = .61$) when included as a covariate, nor did it have any significant interactions ($ps > .34$). As an auxiliary analysis, a regression revealed that, within purity transgressions, both wrongness, $\beta = .59$, $t(52) = 5.91$, $p < .001$, and sadness, $\beta = .28$, $t(52) = 2.78$, $p = .008$, independently predicted the activation of harm, but politics did not, $\beta = .05$, $t(52) = .53$, $p = .60$. In other words, feelings of sadness are tied to perceptions of harm, but this does not account for the link between perceived wrongness and perceived harm; immorality potentiates the perception of victims beyond general negativity.

Across analyses, an implicit measure of social cognition revealed that ostensibly harmless scenarios automatically activate the concept of harm and that this activation cannot be explained by general negativity. Next, we test whether immorality is more closely associated with harm than with general negativity when these concepts are pitted against each other.

Study 3: Harm Implicit Association Test

In this study we tested whether impure actions more robustly activate the concept of harm than the nonmoral negative concept of failure. Switching from the AMP to the Implicit Association Test

(IAT) allowed us to make direct comparisons of implicit associations: If impurity activates harm merely because impurity is negative, we would predict no special association between moral transgressions and perceived harm versus failure. However, if impure scenarios preferentially activate harm (as predicted by dyadic completion), we would predict a stronger association between moral transgressions and harm versus failure.

Method

Participants. Forty-six MTurk participants completed this within-subjects study. Five participants were excluded from analysis for failing at least 20% of the trials, leaving a total of 41 participants (56% liberal, 49% female, $M_{age} = 36$ years, all from the United States).⁶

Procedure. In general, IATs measure implicit association between two pairs of concepts (e.g., family/career, male/female). By looking at the relative response time for categorizing concept-relevant terms under different conceptual pairings, one can make inferences about implicit associations (Greenwald, Nosek, & Banaji, 2003). For example, faster categorization of female terms when female is paired with family (vs. career) is evidence of an implicit association between women and the home. In this study, we contrasted responses regarding category labels (a) harmful vs. failure and (b) immoral vs. not immoral to test whether victimless crimes cluster with harm or relate equally to nonmoral negative terms. We predicted that participants would be faster to categorize the victimless purity violations when immorality was paired with harmful than when immoral was paired with failure.

In this IAT, harmful items were *victim*, *harmful*, and *dangerous*, and failure items were *failure*, *disappointing*, and *lose*. Among items used in measuring the implicit associations with purity violations, immoral items were *sister masturbate*, *animal sex*, and *Bible feces*, and not immoral items were *fold newspaper*, *eat sandwich*, and *ride bus*. Because performance on implicit tasks can be influenced by word length and frequency, categories were matched on these criteria using the MRC Psycholinguistic Database.

It should be noted that we used failure as our negative control instead of “not harmful” for two reasons. First, using “not harmful” would provide the possibility that links between immorality and harm arise merely because of the association between two negations (“not harmful” and “not immoral”). Second, failure is somewhat harmful and so provides a more challenging test of our hypothesis, as immorality has to be associated with harm more than with failure.

Before completing the IAT, participants read all six scenarios, and they were asked to categorize these as moral or immoral. In the key trials of the IAT, participants saw each item (e.g., *victim*) on the middle of the screen while concept labels were paired in the upper left and right of the screen. Participants pressed *E* (for left) or *I* (for right) as quickly as possible to categorize words. Our prediction was that participants would be quickest to categorize

⁵ The negative scenarios also activated harm significantly more than the neutral scenarios did ($p = .03$), but they did not activate wrongness more than the neutral scenarios did ($p = .46$).

⁶ Results remain fundamentally unchanged if these participants are included in analyses.

items when concept labels IMMORAL/HARMFUL and NOT IMMORAL/FAILURE were paired, rather than vice versa. After completing the IAT, participants filled out demographics information.

Results and Discussion

IATs were conducted with Millisecond by Inquisit (Version 4.0.4), and D scores were calculated automatically according to established guidelines (Greenwald et al., 2003). D scores measure the difference between average response time for one combination of pairs (IMMORAL/HARMFUL and NOT IMMORAL/FAILURE) versus the opposite (IMMORAL/FAILURE and NOT IMMORAL/HARMFUL). Positive D scores indicate that the more people associate harmless violations with immorality, the more they linked such violations to harm versus failure. Consistent with dyadic completion, the mean D score was positive ($D = .23$; $SD = .53$), which a one-sample t test revealed was significantly different from zero, $t(40) = 2.71$, $p = .01$. The more people saw impurity acts as immoral (vs. not immoral), the more they appear to link them to harm over failure. The correlation between the D score and politics was not significant, $r(41) = .14$, $p = .39$. This reveals that although conservative participants may see impure violations as relatively more wrong, the relation between wrongness and harm is consistent across the political spectrum. In other words, liberals and conservatives may differ on the moral wrongness of different actions, but this wrongness is consistently linked to perceived harm. Together, these data reveal that moral wrongs are linked to harm more than general negativity, consistent with dyadic completion.

Study 4: Lingering Harm

The previous studies found that moral wrongs activate perceived harm, but it is unknown whether this harm is merely metaphorical (Gutierrez & Giner-Sorolla, 2007) or is actually linked to perceived suffering. To investigate the nature of activated harm, participants read about an ostensibly harmless moral transgression before rating the painfulness of physical injuries in a seemingly unrelated task. Ratings of pain were measured in an unrelated task to decrease the likelihood that people were engaging in effortful justification. Dyadic completion predicts that ratings of physical pain would be correlated with the perceived wrongness of moral transgressions. Of importance, this association should be above and beyond any links between immorality and other negative ratings, as measured by both negative affect and the pleasantness ratings of Chinese characters (similar to Payne et al., 2005).

Method

Participants. One hundred seventy-five participants were recruited through MTurk. Twenty-three participants were excluded for failing the instructional manipulation check, and another 17 participants were excluded for failing to recall the correct story in the manipulation check, leaving 132 participants (46% female, $M_{\text{age}} = 34$ years, 48% liberal, all from the United States).

Procedure. Participants first read one of two scenarios adapted directly from previous research (Graham et al., 2009;

Haidt et al., 1993). Half of participants read about a man burning an American flag on Independence Day to protest the government's involvement in Iraq (Burn Flag). Half of participants read about a man who buys a dead chicken from the grocery store and has sex with it before cooking and eating it (Chicken Sex). After reading these scenarios, participants completed an ostensibly unrelated task.

In a counterbalanced order, participant rated the painfulness of five ambiguous injuries (e.g., stub toe, cut finger, hit head) on 5-point scales ranging from *No Pain at All* (1) to *Extremely Painful* (5), and the un/attractiveness of five Chinese characters on 5-point scales ranging from *Very Unattractive* (1) to *Very Attractive* (5). Participants then judged the moral wrongness of the action from the story, the character of the person, and the punishment deserved, all on 5-point scales ranging from *None/Not at All* (1) to *Extreme(ly)* (5). Moral judgments were measured after the injury/ideogram ratings in order to reduce post hoc justification of moral decisions. Participants then rated their negative affect with a shortened Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) and completed the demographics items as in previous studies.

Index construction. The five pain measures were combined into a pain index ($\alpha = .74$). The five Chinese character ratings were combined and reversed-scored to create a negativity measure ($\alpha = .64$). The three morality questions were combined into a wrongness index ($\alpha = .92$). Finally, participants' ratings of negative items from the PANAS were combined into a single negative affect scale ($\alpha = .89$).

Politics. Because these moral scenarios were adapted from past research on morality and politics, we wanted to ensure that moral judgments were related to political affiliation (Graham et al., 2009). Consistent with this work, there were significant correlations between political affiliation and wrongness for both Burn Flag, $r(71) = .41$, $p < .001$, and Chicken Sex, $r(57) = .34$, $p = .009$.

Results

Correlational analyses revealed that judgments of wrongness were correlated with perceptions of pain for the unrelated events, $r(130) = .24$, $p = .005$, but not with general negativity, $r(130) = -.04$, $p = .67$, and that these correlations significantly differed ($z = 2.29$, $p = .02$). Regression analyses were conducted with pain as the criterion and with wrongness, politics, negative affect, story content, and their respective interaction terms as predictors. It revealed a significant model, $R^2 = .09$, $F(5, 126) = 2.26$, $p = .02$. Out of all predictors, only moral wrongness significantly predicted perceived pain, $\beta = .19$, $t(126) = 2.39$, $p = .02$; it was not predicted by politics ($\beta = .04$), self-reported negative affect ($\beta = -.04$), story content ($\beta = .04$), or general negativity as assessed through attractiveness ratings ($\beta = -.075$). No interactions, including those with politics or story content, were significant. A similar model using wrongness, politics, negative affect, content area, and pain to predict general negativity (attractiveness ratings) did not reveal a significant model, $F(5, 126) = 0.40$, $p = .85$. Consistent with a dyadic template, moral wrongness was tied to perceived pain above and beyond general negativity, for both liberals and conservatives.

Conceptual Replication

As a tougher test of our hypothesis, we attempted to replicate the findings of Study 4 with perceptions of disgust rather than attractiveness as a control. In two separate samples of 100 MTurkers, participants read about flag burning (after attention check exclusions, $N = 80$, 65% female, $M_{\text{age}} = 35$ years, 48% liberal) or eating one's dead dog ($N = 80$; 50% female, $M_{\text{age}} = 33$ years, 53% liberal) and then were asked to rate either the painfulness of injuries or the grossness of potentially disgusting foods (e.g., old milk, sardines). Past research validates measures of grossness in assessing disgust (e.g., Widen & Russell, 2002). As predicted by a dyadic template, moral wrongness correlated with pain for both Burn Flag, $r(41) = .43$, $p < .005$, and Eat Dog, $r(39) = .48$, $p < .005$, but not with grossness of disgusting foods ($r_s < .18$, $p_s > .30$). These results replicate those of Study 4 and are consistent with Study 3, in which immorality was more strongly tied to harm than other negative concepts. This replication is striking, given that many suggest that disgust enjoys a privileged relation with purity judgments (Rozin, Lowery, Imada, & Haidt, 1999; Schnall et al., 2008), but here it was less correlated with immorality than was harm.

Discussion

Study 4 revealed that the cognitive link between perceived immorality and perceived harm is not simply metaphorical; even in objectively harmless scenarios, participants tied judgments of wrongness to perceptions of physical pain. It should be noted that some scenarios did involve some potential for pain (e.g., fire can be physically dangerous); however, general characteristics of scenarios cannot explain correlations between wrongness and perceived pain within each scenario.⁷ These results suggest that although people may see abstract and symbolic harm in moral wrongs, they can also perceive physical suffering. With this design, it is still possible that participants were taking time to think through each of the pain ratings as they related to the moral wrongness of scenarios. To test whether direct suffering is indeed implicitly activated, the following study uses a traditional implicit cognition paradigm: the AMP.

Study 5: The Suffering of Children

In this study, we paired different victimless moral violations with images of potential victims—sad children—in a modified AMP. We predicted that dyadic completion would occur quickly and automatically such that victimless primes would increase ratings of children's suffering relative to neutral primes. To provide a stronger test of our model, we also predicted that victimless transgressions would activate suffering in a manner similar to harmful moral violations (i.e., those with obvious victims). Finally, as a control for affective congruence, we predicted that ratings of a child's boredom would not increase when participants were primed with victimless crimes relative to harmful or neutral primes.

Method

Participants. One hundred and two MTurk participants completed this study on Inquisit (Version 4.0.4). Two participants were

excluded for failing to follow instructions, leaving 100 participants (54% female, $M_{\text{age}} = 34$ years, 51% liberal, all from the United States).

Procedure. Participants first read each of the 12 scenarios used in Study 1, in the three categories of impure (Sister Masturbate; Animal Sex; Defile Corpse; Bible Feces); harmful (Kick Dog; Punch Wife; Stick Pin; Insult Colleague), and neutral (Read Text; Ride Bus; Fold Paper; Eat Toast). To put participants in a moral mindset, we had them categorize the actions as immoral or moral as quickly as they could, as in Study 2. Next, participants completed two AMP blocks—suffering, bored—in random order. Participants first saw a fixation point for 500 ms, followed by the scenario for 250 ms, a blank screen for 150 ms, and a picture of a child's face for 250 ms, followed immediately by a static screen mask. See Figure 4 for a sample image of a child. During the suffering block, participants rated how much suffering the child was experiencing from *No Suffering at All* (1) to *Extreme Suffering* (4). In the bored block participants rated how bored the child was from *Not Bored at All* (1) to *Very Bored* (4). Participants saw each story four times in each of the counterbalanced blocks, for a total of 96 trials.

At the end of the study, participants explicitly rated the suffering and boredom of all photos, providing baseline scores. This was done because, in contrast to the abstract Chinese characters used in Study 2, pictures of children may intrinsically vary on perceived expressions of suffering and boredom. To calculate the amount that scenarios experimentally influenced perceptions of suffering and boredom, we subtracted participants' baseline ratings of photos from their average implicit ratings.⁸ Finally, all participants completed a demographics survey as in previous studies.

Results and Discussion

The child ratings were averaged by scenario type (all $\alpha_s > .80$). A 2 (child rating: suffering, bored) \times 3 (scenario: impure, harmful, neutral) within-subjects ANOVA revealed a significant main effect of scenario, $F(2, 198) = 8.08$, $p < .001$, $\eta^2 = .08$; no main effect of rating, $F(1, 99) = .039$, $p = .85$; and a significant interaction, $F(2, 198) = 6.69$, $p = .002$, $\eta^2 = .06$. See Figure 5 for means. As predicted, participants rated the suffering of the child higher after both impure ($M = .14$, $SD = .53$) and harmful ($M = .09$, $SD = .48$) scenarios than after neutral scenarios ($M = -.11$, $SD = .55$), $p_s < .001$. The harm and impure scenarios did not differ significantly from each other ($p = .12$). If anything, there were higher perceptions of suffering in the impurity condition. Further analysis revealed that impurity scenarios not only increased perceptions of harm relative to neutral scenarios but also increased perceptions of suffering relative to the baseline explicit ratings, $t(99) = 2.73$, $p = .008$. In contrast, ratings of boredom did not vary significantly based on the scenario, $F(2, 198) = 0.85$, $p = .43$. Politics did not correlate significantly with any of the ratings (all $r_s < .13$, $p_s > .21$).

It appears that moral scenarios, whether obviously harmful or ostensibly victimless, increase perceptions of suffering relative to

⁷ Nor can this account for the link between immorality and harm for ostensibly harmless chicken masturbation or dog eating.

⁸ An analysis of covariance is not appropriate here because each measure (suffering and boredom) has its own baseline.



Figure 4. Sample child's face used in Study 5. ©iStockphoto/Crazytang

neutral stories but do not shift ratings of other, nonmoral negative judgments such as boredom. As in previous studies, these data reveal that people implicitly link immorality to suffering. In particular, they reveal that ostensibly harmless wrongs compel people to see the suffering of children.

General Discussion

Across five studies, perceived immorality is linked to perceived harm, even in objectively harmless scenarios, and these perceptions are not simply due to affective congruence (bad = bad). Of importance, perceived harm is unlikely to be the product of effortful rationalization, as revealed by the time pressure manipulation (Study 1), implicit social cognition studies (Studies 2 and 3), unrelated judgments of injuries (Study 4), and perceptions of children's suffering (Study 5). Of course, effortful rationalizations exist and often do focus upon issues of harm (Haidt & Hersh, 2001; Sood & Darley, 2007), but we suggest that these justifications build off of initial and automatic perceptions. Analogously, moral judgments can also involve explicit reasoning (Mercier, 2011; Mercier & Sperber, 2011; Pizarro & Bloom, 2003), but this does not rule out the importance—or psychological persistence—of intuitive perceptions of right or wrong. Dyadic completion simply suggests that harm—like morality—is subjective, and that its perception is cognitively linked to moral judgment through a template that binds together the acts of immoral agents and the suffering of moral patients. What is wrong seems to be harmful. This suggests that harm perceived by opponents of ostensibly harmless acts such as masturbation, gay marriage, or recreational drug use does not reflect aberrant mental tendencies or labored justification but instead the general tendency of dyadic completion.

Scope of the Current Research

It is important to acknowledge the scope of these studies. In revealing a persistent link between immorality and the perception of harmed victims, these studies did not demonstrate that various transgressions are wrong because they are harmful. Certainly, harm is causally—and consistently—linked to immorality (e.g., Turiel, 1983) and increases in perceived threat does induce moralization (Eibach, Libby, & Ehrlinger, 2009; Sheikh, Ginges, Co-

man, & Atran, 2012; van Leeuwen & Park, 2009), but it is not the only route to moral judgment, which can be triggered by noxious smells (Schnall et al., 2008), aversive environments (Eskine, Kacinik, & Prinz, 2011), hypnosis (Wheatley & Haidt, 2005), violations of diverse norms (Monroe et al., 2012), or any combination of the above (Rottman & Kelemen, 2012). No matter how a moral judgment is triggered, however, we suggest that a dyadic template compels perceptions of moral patients through top-down influences. Similar processes occur in the realm of stereotypes: Judgments of a person's race can proceed through various routes, but, once triggered, racial stereotypes implicitly shape perceptions of that person (Macrae & Bodenhausen, 2000). These stereotypes not only are implicit and automatic but also guide behavior and subsequent judgments (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). The moral dyad could thus be thought of as a "stereotype" of moral situations, a social cognitive template that both shapes and is shaped by experience in a dynamic feedback process.

We also acknowledge that these studies focused primarily upon Western participants. Research suggests that both moral and nonmoral judgments made by WEIRD (White, educated, industrialized, rich, democratic; Henrich, Heine, & Norenzayan, 2010) participants may not generalize to all people, and so it would be useful to replicate these studies cross-culturally. However, research also suggests that American conservatives are a distinct cultural group with diverse moral concerns (Graham et al., 2009), and the current studies did include conservatives. Moreover, the original moral dumbfounding study—which has been used to argue against the persistence of harm—also used a WEIRD sample of 30 college students (Haidt et al., 2000).

Although we do find significant correlations between politics and explicit ratings of immorality, our studies reveal that these cultural differences might not be as deep as past theorizing suggests. Our implicit measures did not reveal the political moral divide found in research using explicit ratings (Graham et al., 2009). Instead, we found substantial similarity between liberals and conservatives, consistent with other research using implicit (Wright & Baril, 2011) and behavioral measures (Skitka & Bauman, 2008) and studies examining individual differences (Frimer, Biesanz, Walker, & MacKinlay, 2013). More important, our stud-

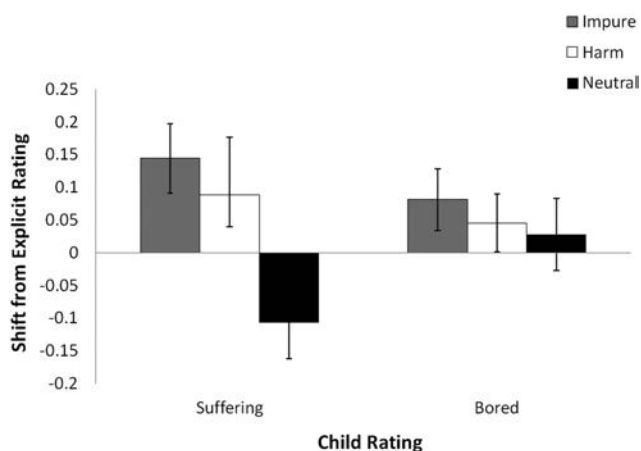


Figure 5. Average rating of children's suffering or boredom based on the story prime (Study 5). Error bars represent standard error.

ies found that the process of dyadic completion occurs similarly across liberals and conservatives. Even if moral judgments differ across the political spectrum, if an act is seen as wrong, then people are potentiated to see a moral patient suffering harm.

This research examined a wide variety of moral violations, but it could not examine all potential moral violations. It is therefore possible that some moral violations remain unlinked to perceived harm.⁹ Nevertheless, these studies provide a conservative test of dyadic completion by using moral scenarios constructed by others (Haidt et al., 1993) that were “carefully written to be harmless” (Haidt et al., 2000, p. 6). We welcome future research to evaluate the generalizability of dyadic completion.

One open question in dyadic completion concerns the identity of perceived victims: Who exactly is victimized by masturbation? In our studies, participants were not asked to explicitly identify victims because such identification likely involves effortful reasoning—the very process we tried to rule out. Nevertheless, it is an interesting question, and so we conducted a follow-up study in which 85 MTurk participants (all American, 32% female, $M_{age} = 34$ years, 54% liberal) identified the victims they perceived in our four ostensibly harmless impurity scenarios. In *Sister Masturbate*, participants who perceived victims perceived harm to the sister’s spirit/memory (45%), the perpetrator him/herself (24%), or the family and society (10%).¹⁰ In *Bible Feces*, people perceived harm to believers (42%), God (25%), the perpetrator (13%), or the Bible (4%). In *Defile Corpse*, people perceived harm to the corpse/memory/soul of the dead person (42%), the perpetrator him/herself (20%), and the family (39%). Finally, in *Animal Sex*, people perceived harm to the animals (40%), the perpetrator (35%), or society (5%). Integrating across these responses, perceived victims fell roughly into one of three categories: harm to the self, another person/soul, or society in general. This variety of victims nicely matches the model of moral motives provided by Janoff-Bulman and Carnes (2013), who suggested that different moral motives of approach (help) and avoidance (harm) are expressed differently across the self, a specific other, and society at large.

This alignment suggests that dyadic morality is consistent with some accounts of moral pluralism, in that common psychological processes (i.e., dyadic completion) can give rise to descriptively different judgments (i.e., kinds of victims). Of importance, the presence of different victims does not require different psychological *mechanisms*. Common psychological processes can give rise to different religions (Boyer, 2001; Norenzayan & Shariff, 2008), different emotions (Gray & Wegner, 2011a; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), and different stereotypes (Bodenhausen, 1990). For example, stereotypes can have diverse contents—people perceive Whites as unathletic (Stone, Perry, & Darley, 1997), Blacks as unmotivated (Devine, 1989), and Asians as unfeeling (Bain, Park, Kwok, & Haslam, 2009)—but each nevertheless involves a top-down influence of conceptual prototypes on perceptions of individuals.

The concision between common processes and descriptive differences can be seen in agentic dyadic completion. When people blame agents for wrongdoing, they often point to God (Gray & Wegner, 2010a), the government (Kay, Gaucher, Napier, Callan, & Laurin, 2008), and even calculating swine (Oldridge, 2004), depending upon which is most cognitively accessible. We suggest a similar process with patientic dyadic completion, such that cognitive accessibility would predict the exact victim identified.

Such accessibility would undoubtedly be influenced by individual differences—it is more likely for the religious than atheists to perceive spiritual suffering (e.g., the tainting of the soul). Of course, there are certainly better and worse candidates for victims, and dyadic morality predicts that those with ample experience and little agency make ideal moral patients because of moral typecasting, or the tendency to divide other minds into agents or patients (Gray et al., 2007; Gray & Wegner, 2009). This is why people frequently perceive ostensibly victimless wrongs such as homosexuality as harming children (high experience, low agency; Bryant, 1977) rather than corporations (low experience, high agency; Knobe & Prinz, 2008).

Implications and Extensions

Dyadic completion is a relatively modest claim—perceived immorality is intuitively linked to perceived harm—but we suggest that it has broad implications for moral psychology. First, it emphasizes the importance of perceived harm over objective harm. Although various scenarios may exclude objective harm, subjective harm may persist at an intuitive level. In the terms of Gendler (2008), beliefs—explicit knowledge—can be separated from aliefs, which are affectively laden intuitions. The importance of intuitions is consistent with the social intuitionist model (Haidt, 2001) and heuristics in moral and nonmoral decision making (Gigerenzer, 2008; Gilovich, Griffin, & Kahneman, 2002; Sunstein, 2005), which both highlight how intuitive judgments can diverge from objective rationality. If people consistently make inaccurate judgments about probability and value in the face of dispassionate data (Tversky & Kahneman, 1974), what hope is there for objectivity in emotionally bound judgments of suffering? As judgments of harm hinge on ambiguous mind perception, it may be as impossible to objectively gauge harm as it is to objectively gauge immorality, beauty, pornography, or art. Each of these things depends on the idiosyncratic eye of the perceiver and therefore defies objective definitions. What objective definition of harm can reconcile historical and political contradictions such as viewing genocide as a “solution” and birth control as “murder”?

Second, the perceived nature of harm highlights the importance of considering the perceiver. Research suggests that liberals may have a narrower moral scope than conservatives do, and fail to see some acts of disloyalty and impurity as immoral (Graham et al., 2009; but see Frimer et al., 2013; Janoff-Bulman & Carnes, 2013). The liberal orientation of social psychology likely prevented researchers from seeing these violations as legitimately moral (Inbar & Lammers, 2012), and we suggest that the same liberal bias has prevented researchers from accepting these acts as legitimately harmful. Liberal researchers may believe impurity scenarios to be “harmless offenses” (Haidt et al., 1993, p. 613), but this does not mean that conservative participants share this opinion. Gay-rights opponents frequently cite the harm homosexuality does to children (*Hollingsworth v. Perry*, 2013), and though such statements may

⁹ The scientific difficulty of induction—generalizing from specific examples to general principles—has been long recognized (Hume, 1748/2003; Popper, 1959). Even though all observed ravens are black, one cannot be sure that a white raven does not lurk somewhere (Hempel, 1945).

¹⁰ Percentages do not add to 100%, because some participants declined to identify specific victims.

be dismissed by others as mere rhetoric, we suggest that these victimless acts are genuinely perceived as harmful within some moral communities. Just as with moral judgment, understanding perceptions of harm requires cultural sensitivity—"harmless wrongs" seem to exist only in the minds of liberals who characterize conservative morality (e.g., Haidt, 2012), and not in the minds of conservatives themselves.

Third, the persistence of perceived victims challenges theories that postulate moral domains completely independent of harm. Recent research has emphasized cultural differences across moral content, dividing the moral sphere into various moral "foundations" of harm, fairness, loyalty, authority, and purity (Graham et al., 2013). These cultural differences are hypothesized to stem from domain specific moral mechanisms (Haidt & Joseph, 2007), each of which independently processes moral content. Although descriptive cultural differences between moral content are both well documented and important (Ditto & Koleva, 2011), they need not imply distinct cognitive modules. There may be cultural differences between the United States and India in morality (Shweder, Mahapatra, & Miller, 1987), but there are also differences in food and fashion preferences. The latter do not imply distinct modules for pizza and curry or jeans and saris.

Empirically, factor analyses reveal substantial overlap between even descriptive content, as loyalty, authority, and purity are highly correlated with each other (mean $r = .80$), as are the factors of harm and fairness ($r = .72$; Graham et al., 2011). Many of these between-factor correlations are actually higher than the test-retest reliability (i.e., within-factor correlation) of content areas (e.g., loyalty $r = .71$; fairness $r = .69$; Graham et al., 2011, p. 371), explaining why extreme power ($N = 35,000$) is needed to resolve moral judgments into a five-factor model. Even a two-factor descriptive description is called into question by recent discoveries of sampling bias and confounds across scenario domains (Gray & Keeney, 2013). Harm and purity may load onto different factors simply because purity violations are weirder.

Even if these descriptive differences are taken at face value, the current studies suggest that perceived harm is a potential source of cognitive overlap between different moral content. Whether violations involve affronts to patriotism (flag burning), purity (masturbation), or sanctity (desecrating a Bible), people perceive them to cause harm. This perceived harm may be post hoc but appears

not to need conscious elaboration. This possibility lends support to the idea that diverse concerns—even if initially separable—are still viewed through a dyadic template (Gray, Waytz, & Young, 2012; Gray, Young, & Waytz, 2012). This interpretation is consistent with that of researchers who emphasize the importance of domain general moral processes including intention, causation, and coalition building (Cushman & Young, 2011; DeScioli & Kurzban, 2013; Shenhav & Greene, 2010; Strickland et al., 2012).

Fourth, dyadic completion provides a cautionary note about the use of moral dilemmas that pit deontology (immoral acts) against utilitarianism (good outcomes; see also Gray & Schein, 2012; Liu & Ditto, 2012). Moral dilemmas have helped to reveal a number of important factors that influence of moral judgment—including physical force, omissions versus commission and the directness of harm (Cushman & Greene, 2012; Cushman et al., 2006; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001)—but may not represent typical moral scenarios. Moral dilemmas force participants to choose between immoral actions or harmful consequences (Philippa, 1967; Thomson, 1976). For example, do you commit murder to save more lives? In the framework of dyadic morality, this choice translates to selecting either immoral agents or suffering moral patients. A dyadic template suggests that this choice is akin to asking people whether marriage is about husbands or wives. The answer, of course, is both.

As the current studies demonstrate, moral cognition is likely characterized not by conflict between actions/agents and outcomes/patients but by their mutual activation. The dyadic template binds together immoral actions with bad outcomes, explaining why people almost never say that evil acts are beneficial and harmful acts are morally permissible (Kahan, 2007; Liu & Ditto, 2012). Indeed, we suggest that moral dilemmas may be better thought of as moral *paradoxes* because they are *anti-dyadic*, stymying the cognitive processes that attempt to reconcile agents and patients through dyadic completion (see Figure 6). This paradoxical nature explains why dilemmas are so emotionally evocative (Greene et al., 2001); holding two incongruent attitudes is the hallmark of dissonance (Festinger & Carlsmith, 1959). Atypical paradoxes are fun to consider but may give a misleading picture of moral cognition (for a detailed discussion, see Gray & Schein, 2012). In particular, the separation (and conflict) of affect and cognition posited by dual-process models (Greene, Morelli, Lowenberg,

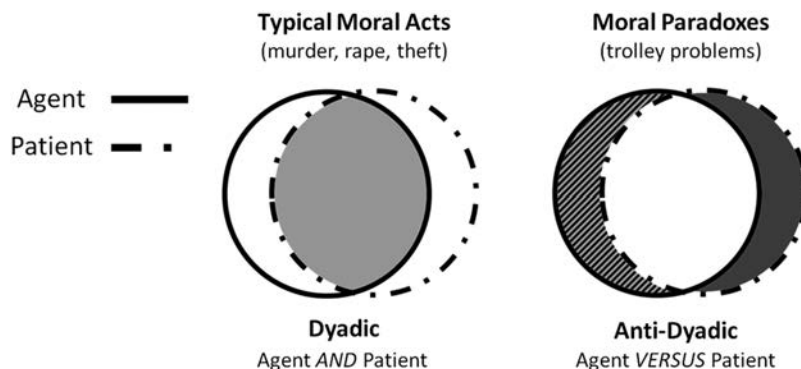


Figure 6. Typical moral acts are dyadic, containing both an agent and a patient. Scenarios used to test many claims about moral judgment cognition are anti-dyadic (e.g., trolley scenarios), forcing a choice between agents or patients.

Nystrom, & Cohen, 2008; Greene, Nystrom, Engell, Darley, & Cohen, 2004) may not persist for moral judgment of typical scenarios, in which affect and cognition appear to both overlap (Pessoa, 2008) and work together (Haidt, 2001).

Conclusion

Across diverse psychological arenas, the top-down influence of cognitive working models is uncontroversial: Visual experience is shaped by gestalt principles, social experience is shaped by stereotypes, and nonmoral decision making is shaped by heuristics. In this paper, we suggest that moral judgments are also shaped by the top-down influence of the moral dyad. Dyadic completion helps explain not only belief in God (the ultimate moral agent; Gray & Wegner, 2010a) and the labile nature of causal judgments in moral contexts (Alicke, 1992) but also the enduring perceived presence of harm in immorality. Moral judgments may vary widely across cultures and may involve affronts to oneself, one's country, and one's God, but each of these appears to trigger implicit perceptions of harm.

This link is perhaps best exemplified by Anita Bryant, the once successful country singer who titled her autobiography *The Anita Bryant Story: The Survival of Our Nation's Families and the Threat of Militant Homosexuality*. Bryant was so deeply convinced of the harm of gay rights that she sacrificed her career, her fortune, her house, and her husband trying to fight them. Homosexuality may seem harmless to many, but for her—and millions like her—the link between sin and suffering is self-evident. These examples and the current studies suggest that the very idea of “victimless wrongs” may be a psychological impossibility, a moral paradox that defies a dyadic template. For those who see these acts as immoral, the link between morality and harm is not so easily severed. In the spotlight of moral judgment, harm is an ever-present shadow. Even without objective substance, people cannot help but see its darkness.

References

- Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, *63*, 368–378. doi:10.1037/0022-3514.63.3.368
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*, 556–574. doi:10.1037/0033-2909.126.4.556
- Ames, D. L., & Fiske, S. T. (2013). Intentional harms are worse, even when they're not. *Psychological Science*, *24*, 1755–1762. doi:10.1177/0956797613480507
- Appiah, A. (2008). *Experiments in ethics*. Cambridge, MA: Harvard University Press.
- Bain, P., Park, J., Kwok, C., & Haslam, N. (2009). Attributing human uniqueness and human nature to cultural groups: Distinct forms of subtle dehumanization. *Group Processes & Intergroup Relations*, *12*, 789–805. doi:10.1177/1368430209340415
- 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. doi:10.1037/0022-3514.71.2.364
- Bastian, B., Costello, K., Loughnan, S., & Hodson, G. (2012). When closing the human–animal divide expands moral concern: The importance of framing. *Social Psychological & Personality Science*, *3*, 421–429. doi:10.1177/1948550611425106
- Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, *1*, 319–322. doi:10.1111/j.1467-9280.1990.tb00226.x
- Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York, NY: Basic Books.
- Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*, 237–273. doi:10.1016/0010-0277(83)90006-9
- Bryant, A. (1977). *The Anita Bryant story: The survival of our nation's families and the threat of militant homosexuality*. Grand Rapids, MI: Revell.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3–5. doi:10.1177/1745691610393980
- Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology*, *42*, 437–451.
- Castano, E., & Giner-Sorolla, R. (2006). Not quite human: Infrahumanization in response to collective responsibility for intergroup killing. *Journal of Personality and Social Psychology*, *90*, 804–818. doi:10.1037/0022-3514.90.5.804
- Chalmers, D. J. (1997). *The conscious mind: In search of a fundamental theory*. New York, NY: Oxford University Press.
- Craik, K. J. W. (1967). *The nature of explanation*. Cambridge, United Kingdom: Cambridge University Press.
- 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., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience*, *7*, 269–279. doi:10.1080/17470919.2011.614000
- 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
- Davis, M. H. (1996). *Empathy: A social psychological approach*. Boulder, CO: Westview Press.
- 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., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, *20*, 1053–1080. doi:10.1017/S0954579408000503
- DeScioli, P. (2008). *Investigations into the problems of moral cognition* (Doctoral dissertation, University of Pennsylvania). Retrieved from <http://repository.upenn.edu/dissertations/AAI3309424/>
- DeScioli, P., Gilbert, S., & Kurzban, R. (2012). Indelible victims and persistent punishers in moral cognition. *Psychological Inquiry*, *23*, 143–149. doi:10.1080/1047840X.2012.666199
- DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*, 477–496. doi:10.1037/a0029065
- Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on “automatic evaluations” as a function of task characteristics of the measure. *Journal of Experimental Social Psychology*, *45*, 101–114. doi:10.1016/j.jesp.2008.09.001
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18. doi:10.1037/0022-3514.56.1.5
- Ditto, P. H., & Koleva, S. P. (2011). Moral empathy gaps and the American culture war. *Emotion Review*, *3*, 331–332. doi:10.1177/1754073911402393

- Ditto, P. H., & 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.
- 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.
- Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*, 327–339. doi:10.1037/0022-3514.87.3.327
- Epley, N., & Waytz, A. (2009). Mind perception. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (5th ed., pp. 498–541). New York, NY: Wiley.
- Eskine, K. J., Kacirik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science, 22*, 295–299. doi:10.1177/0956797611398497
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*, 203–210. doi:10.1037/h0041593
- Frimer, J. A., Biesanz, J. C., Walker, L. J., & MacKinlay, C. W. (2013). Liberals and conservatives rely on common moral foundations when making moral judgments about influential people. *Journal of Personality and Social Psychology, 104*, 1040–1059. doi:10.1037/a0032277
- Gardner, W. L., & Knowles, M. L. (2008). Love makes you real: Favorite television characters are perceived as “real” in a social facilitation paradigm. *Social Cognition, 26*, 156–168. doi:10.1521/soco.2008.26.2.156
- Gawronski, B., & Ye, Y. (2014). What drives priming effects in the affect misattribution procedure? *Personality and Social Psychology Bulletin, 40*, 3–15. doi:10.1177/0146167213502548
- Gendler, T. S. (2008). Alief and belief. *Journal of Philosophy, 105*, 634–663.
- Gigerenzer, G. (2008). Moral intuition = fast and frugal heuristics? In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality: Intuition and diversity* (pp. 1–26). Cambridge, MA: MIT Press.
- Gilbert, D. T., & Gill, M. J. (2000). The momentary realist. *Psychological Science, 11*, 394–398. doi:10.1111/1467-9280.00276
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, United Kingdom: Cambridge University Press.
- 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
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*, 213–224. doi:10.1002/bdm.1753
- 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
- Gorney, C. (2011). Too young to wed: The secret world of child brides. *National Geographic, 219*(6), 78–99.
- 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., 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., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366–385. doi:10.1037/a0021847
- Gray, H. M., Gray, K., & Wegner, D. M. (2007, February 2). Dimensions of mind perception. *Science, 315*, 619. doi:10.1126/science.1134475
- Gray, K. (2012). The power of good intentions: Perceived benevolence soothes pain, increases pleasure, and improves taste. *Social Psychological & Personality Science, 3*, 639–645. doi:10.1177/1948550611433470
- 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, USA, 108*, 477–479. doi:10.1073/pnas.1015493108
- Gray, K., & Keeney, J. (2013). *Domains or dimensions: Severity and atypicality explain the role of intention across moral content areas*. Manuscript in preparation, University of North Carolina.
- Gray, K., Knickman, T. A., & Wegner, D. M. (2011). More dead than dead: Perceptions of persons in the persistent vegetative state. *Cognition, 121*, 275–280. doi:10.1016/j.cognition.2011.06.014
- Gray, K., & 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*, 405–423. doi:10.1007/s13164-012-0112-5
- Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry, 23*, 206–215. doi:10.1080/1047840X.2012.686247
- Gray, K., & Wegner, D. M. (2008). The sting of intentional pain. *Psychological Science, 19*, 1260–1262. doi:10.1111/j.1467-9280.2008.02208.x
- Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology, 96*, 505–520. doi:10.1037/a0013748
- Gray, K., & Wegner, D. M. (2010a). Blaming God for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review, 14*, 7–16. doi:10.1177/1088868309350299
- Gray, K., & Wegner, D. M. (2010b). Torture and judgments of guilt. *Journal of Experimental Social Psychology, 46*, 233–235. doi:10.1016/j.jesp.2009.10.003
- Gray, K., & Wegner, D. M. (2011a). Dimensions of moral emotions. *Emotion Review, 3*, 258–260. doi:10.1177/1754073911402388
- Gray, K., & Wegner, D. M. (2011b). To escape blame, don't be a hero—Be a victim. *Journal of Experimental Social Psychology, 47*, 516–519. doi:10.1016/j.jesp.2010.12.012
- Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition, 125*, 125–130. doi:10.1016/j.cognition.2012.06.007
- Gray, K., & Wegner, D. M. (2013). Six guidelines for interesting research. *Perspectives on Psychological Science, 8*, 549–553. doi:10.1177/1745691613497967
- 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
- Greenberg, K. J., & Dratel, J. L. (Eds.). (2005). *The torture papers: The road to Abu Ghraib*. Cambridge, United Kingdom: Cambridge University Press.
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences, 359*, 1775–1785. doi:10.1098/rstb.2004.1546
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*, 1144–1154. doi:10.1016/j.cognition.2007.11.004
- Greene, J. D., 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. doi:10.1016/j.neuron.2004.09.027

- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001, September 14). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108. doi:10.1126/science.1062872
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216. doi:10.1037/0022-3514.85.2.197
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*, 17–41. doi:10.1037/a0015575
- Guglielmo, S., Monroe, A. E., & Malle, B. F. (2009). At the heart of morality lies folk psychology. *Inquiry*, *52*, 449–466. doi:10.1080/00201740903302600
- 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
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834. doi:10.1037/0033-295X.108.4.814
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Pantheon.
- Haidt, J., & Bjorklund, F. (2007). Social intuitionists answer six questions about morality. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality: Intuition and diversity* (pp. 181–217). Cambridge, MA: MIT Press.
- Haidt, J., Bjorklund, F., & Murphy, S. (2000). *Moral dumbfounding: When intuition finds no reason*. Unpublished manuscript, University of Virginia.
- Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*, 191–221. doi:10.1111/j.1559-1816.2001.tb02489.x
- 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. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind* (Vol. 3, pp. 367–391). New York, NY: Oxford University Press.
- 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. doi:10.1037/0022-3514.65.4.613
- Harm. (n.d.). In *Merriam-Webster's online dictionary*. Retrieved from <http://www.merriam-webster.com/dictionary/harm?show=0&t=1391643541>
- Hart, H. L. A., & Honoré, T. (1985). *Causation in the law* (2nd ed.). New York, NY: Oxford University Press.
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252–264. doi:10.1207/s15327957pspr1003_4
- Haslam, N., Loughnan, S., Kashima, Y., & Bain, P. (2008). Attributing and denying humanness to others. *European Review of Social Psychology*, *19*, 55–85. doi:10.1080/10463280801981645
- Hauser, M., Young, L., & Cushman, F. (2007). Reviving Rawls's linguistic analogy: Operative principles and the causal structure of moral actions. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality: Intuition and diversity* (pp. 107–144). Cambridge, MA: MIT Press.
- Hempel, C. G. (1945). Studies in the logic of confirmation (I.). *Mind*, *54*, 1–26. doi:10.1093/mind/LIV.213.1
- Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–83. doi:10.1017/S0140525X0999152X
- Hollingsworth v. Perry. (2013, March 26). Retrieved from http://www.supremecourt.gov/oral_arguments/argument_transcripts/12-144.pdf
- Hume, D. (2003). *An enquiry concerning human understanding*. Retrieved from <http://www.gutenberg.org/ebooks/9662> (Original work published 1748)
- Imhoff, R., Schmidt, A. F., Bernhardt, J., Dierksmeier, A., & Banse, R. (2011). An inkblot for sexual preference: A semantic variant of the affect misattribution procedure. *Cognition & Emotion*, *25*, 676–690. doi:10.1080/02699931.2010.508260
- Inbar, Y., & Lammers, J. (2012). Political diversity in social and personality psychology. *Perspectives on Psychological Science*, *7*, 496–503. doi:10.1177/1745691612448792
- Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin*, *38*, 52–62. doi:10.1177/0146167211430232
- Jacobson, D. (2012). Moral dumbfounding and moral stupefaction. In M. Timmons (Ed.), *Oxford studies in normative ethics* (Vol. 2, pp. 289–316). New York, NY: Oxford University Press.
- Jahoda, G. (1998). *Images of savages: Ancient roots of modern prejudice in western culture*. New York, NY: Routledge.
- James, W. (1890). *The principles of psychology*. Cambridge, MA: Harvard University Press.
- 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
- Kahan, D. M. (2007). The cognitively illiberal state. *Stanford Law Review*, *60*, 115–154.
- Kapelner, A., & Chandler, D. (2010, October.). *Preventing satisficing in online surveys: A "kapcha" to ensure higher quality data*. Paper presented at CrowdConf, San Francisco, CA. Retrieved from <http://www.danachandler.com/files/kapcha.pdf>
- Karpman, S. (1968). Fairy tales and script drama analysis. *Transactional Analysis Bulletin*, *7*, 39–43.
- Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology*, *95*, 18–35. doi:10.1037/0022-3514.95.1.18
- Kellogg, J. H. (1890). *Plain facts for old and young: Embracing the natural history and hygiene of organic life*. Retrieved from <http://archive.org/details/plainfactsforold00kell>
- Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, *63*, 190–194. doi:10.1093/analys/63.3.190
- Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences*, *7*, 67–83. doi:10.1007/s11097-007-9066-y
- 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., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, *19*, 448–468. doi:10.1016/0022-1031(83)90022-7
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*, 121–143. doi:10.1017/S0140525X11000446
- Liu, B. S., & Ditto, P. H. (2013). What dilemma? Moral evaluation shapes factual belief. *Social Psychological & Personality Science*, *4*, 316–323. doi:10.1177/1948550612456045
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*, 93–120. doi:10.1146/annurev.psych.51.1.93
- Mercier, H. (2011). What good is moral reasoning? *Mind & Society*, *10*, 131–148. doi:10.1007/s11299-011-0085-6

- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57–74. doi:10.1017/S0140525X10000968
- 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
- Mill, J. S. (2008). Utilitarianism. In M. Warnock (Ed.), *Utilitarianism and On Liberty* (2nd ed., pp. 181–235). doi:10.1002/9780470776018.ch4 (Original work published 1863)
- 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
- Müller-Lyer, F. C. (1889). Optische urteilstauschungen [Optical illusions judgment]. *Archiv für Anatomie und Physiologie, Physiologische Abteilung*, *2*(Suppl.), 263–270.
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, *84*, 221–236. doi:10.1016/S0010-0277(02)00048-3
- Norenzayan, A., & Shariff, A. F. (2008, October 3). The origin and evolution of religious prosociality. *Science*, *322*, 58–62. doi:10.1126/science.1158757
- 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.
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*, 867–872. doi:10.1016/j.jesp.2009.03.009
- 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
- Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, *36*, 163–177. doi:10.1111/j.1551-6709.2011.01210.x
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293. doi:10.1037/0022-3514.89.3.277
- Payne, J. W., & Bettman, J. R. (2004). Walking with the scarecrow: The information-processing approach to decision research. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 110–132). Oxford, United Kingdom: Blackwell.
- Peery, D., & Bodenhausen, G. V. (2008). Black + White = Black: Hypodescent in reflexive categorization of racially ambiguous faces. *Psychological Science*, *19*, 973–977. doi:10.1111/j.1467-9280.2008.02185.x
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*, 148–158. doi:10.1038/nrn2317
- Philippa, F. (1967). The problem of abortion and the doctrine of double effect. *Virtues and Vices*, 19–32.
- Piaget, J. (1932). *The moral judgment of the child*. New York, NY: Harcourt Brace.
- Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: A comment on Haidt (2001). *Psychological Review*, *110*, 193–196. doi:10.1037/0033-295X.110.1.193
- Popper, K. R. (1959). *The logic of scientific inquiry*. London, England: Hutchinson.
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*, 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
- 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
- Roepke, D. (2010). *How risky is it, really? Why our fears don't always match the facts*. New York, NY: McGraw-Hill Professional.
- Rottman, J., & Kelemen, D. (2012). Aliens behaving badly: Children's acquisition of novel purity-based morals. *Cognition*, *124*, 356–360. doi:10.1016/j.cognition.2012.06.001
- 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
- Sava, F. A., MaricuToiu, L. P., Rusu, S., Macinga, I., Vîrgă, D., Cheng, C. M., & Payne, B. K. (2012). An inkblot for the implicit assessment of personality: The semantic misattribution procedure. *European Journal of Personality*, *26*, 613–628. doi:10.1002/per.1861
- Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, *34*, 1096–1109. doi:10.1177/0146167208317771
- Sheikh, H., Ginges, J., Coman, A., & Atran, S. (2012). Religion, group threat and sacred values. *Judgment and Decision Making*, *7*, 110–118.
- Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*, 667–677. doi:10.1016/j.neuron.2010.07.020
- Shukla, G. D., Sahu, S. C., Tripathi, R. P., & Gupta, D. K. (1982). Phantom limb: A phenomenological study. *British Journal of Psychiatry*, *141*, 54–58. doi:10.1192/bjp.141.1.54
- 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.
- Skitka, L. J., & Bauman, C. W. (2008). Moral conviction and political engagement. *Political Psychology*, *29*, 29–54. doi:10.1111/j.1467-9221.2007.00611.x
- Skitka, L. J., & Sargis, E. G. (2006). The Internet as psychological laboratory. *Annual Review of Psychology*, *57*, 529–555. doi:10.1146/annurev.psych.57.102904.190048
- Smith, A. (1982). *The theory of moral sentiments* (D. D. Raphael & A. L. Macfie, Eds.). Indianapolis, IN: Liberty Fund. (Original work published 1759)
- Sood, A. M., & Darley, J. M. (2007, November). *The plasticity of harm in the service of punishment goals: An experimental demonstration*. Paper presented at the Conference on Empirical Legal Studies, New Haven, CT.
- Sousa, P., Holbrook, C., & Piazza, J. (2009). The morality of harm. *Cognition*, *113*, 80–92. doi:10.1016/j.cognition.2009.06.015
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, *27*, 76–105. doi:10.1016/0022-1031(91)90011-T
- Stone, J., Perry, W., & Darley, J. M. (1997). “White men can't jump”: Evidence for the perceptual confirmation of racial stereotypes following a basketball game. *Basic and Applied Social Psychology*, *19*, 291–306. doi:10.1207/s15324834basps1903_2
- 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
- Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, *28*, 531–573. doi:10.1017/S0140525X05000099
- Svenson, O., & Maule, A. J. (Eds.). (1993). *Time pressure and stress in human judgment and decision making*. New York, NY: Plenum Press.
- Tam, K.-P., Lee, S.-L., & Chao, M. M. (2013). Saving Mr. Nature: Anthropomorphism enhances connectedness to and protectiveness toward nature. *Journal of Experimental Social Psychology*, *49*, 514–521. doi:10.1016/j.jesp.2013.02.001

- Tannenbaum, D., Uhlmann, E. L., & Diermeier, D. (2011). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology, 47*, 1249–1254. doi:10.1016/j.jesp.2011.05.010
- Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist, 59*, 204–217. doi:10.5840/monist197659224
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, United Kingdom: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131. doi:10.1126/science.185.4157.1124
- 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
- Ward, A. F., Olsen, A. S., & Wegner, D. M. (2013). The harm-made mind observing victimization augments attribution of minds to vegetative patients, robots, and the dead. *Psychological Science, 24*, 1437–1445. doi:10.1177/0956797612472343
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. doi:10.1037/0022-3514.54.6.1063
- Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who sees human? *Perspectives on Psychological Science, 5*, 219–232. doi:10.1177/1745691610369336
- Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social cognition unbound. *Current Directions in Psychological Science, 19*, 58–62. doi:10.1177/0963721409359302
- Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences, 14*, 383–388. doi:10.1016/j.tics.2010.05.006
- Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of conduct of social conduct*. New York, NY: Guilford Press.
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*, 780–784. doi:10.1111/j.1467-9280.2005.01614.x
- Widen, S. C., & Russell, J. A. (2002). Gender and preschoolers' perception of emotion. *Merrill-Palmer Quarterly, 48*, 248–262. doi:10.1353/mpq.2002.0013
- Wright, J. C., & Baril, G. (2011). The role of cognitive resources in determining our moral intuitions: Are we all liberals at heart? *Journal of Experimental Social Psychology, 47*, 1007–1012. doi:10.1016/j.jesp.2011.03.014

Appendix

Scenarios Used in Studies 1, 2, 3, and 4

Purity Scenarios

Sister Masturbate: A man masturbates to a picture of his dead sister.

Animal Sex: A man watches videos of animals' copulation to become sexually aroused.

Defile Corpse: A man has sex with a corpse.

Bible Feces: A man rubs feces over the Bible.

Harm Scenarios

Kick Dog: Kicking a dog in the head, hard.

Punch Wife: Punching your wife.

Stick Pin: Sticking a pin into the palm of an adult you don't know.

Insult Colleague: Making cruel remarks to an overweight colleague about her appearance.

Negative Nonmoral Scenarios

Lose Teddy: A little girl loses her favorite teddy bear.

Fail Exam: A student fails an exam in an important class and will fail the semester.

Partner Leave: After a 20-year marriage, a woman leaves her husband.

Cat Missing: A little boy loses his beloved kitty cat.

Neutral Scenarios

Read Text: A student reads an article for one of her classes.

Ride Bus: A woman rides a bus to work.

Fold Paper: A man folds a letter to place in an envelope.

Eat Toast: A student eats a piece of toast for breakfast.

Received October 26, 2011

Revision received January 27, 2014

Accepted January 28, 2014 ■