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

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 In an oft-cited but unpublished study (Haidt, Bjorklund, & Murphy, 2000), participants judged as immoral some scenarios engineered to be objectively harmless (e.g., consensual incest).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...
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 harmfulness 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 “objectively” 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; Oosofsky, 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 (Haidt & Bjorklund, 2007). For example, people mete out punishments 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 victim (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 & Honore, 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 \([AP]\), 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 patients \([AP]\), 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 (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

---

1 Figure 1. The Kanizsa triangle: an example of automatic visual completion.
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_{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 immoral-}

Results and Discussion

To examine the role of effortful reasoning in perceived victimhood, we conducted a 3 (scenario: impure, harm, neutral) × 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, η^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, η^2 = .13.

Of importance, these main effects were qualified by a significant interaction, F(2, 160) = 15.43, p < .001, η^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 × 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, β = .25, t(78) = 2.43, p = .02, and time pressure, β = −.35, t(78) = −3.55, p = .001, but not politics, β = .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 (α = .69). Analysis reported here exclude this scenario; however, results are nearly identical with it included.
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_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 (α > .74), neutral (α > .76), and nonmoral negative (α > .46).4 A 3 (rating: harmful, sad, wrong) × 3 (scenario: impure, negative control, neutral control) within-subject ANOVA revealed a significant main effect of rating, F(2, 110) = 11.32, p < .001, η^2 = .17, and scenario, F(2, 110) = 51.74, p < .001, η^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, η^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 (α > .65). Analysis reported here excludes Partner Leave; however, results are nearly identical with this scenario included.
Moral transgressions: 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).  

**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, & Bañaji, 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 (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.

**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.

**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).

**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, & Bañaji, 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
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 word, 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_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 Extremely (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 shortenned 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 (α = .74). The five Chinese character ratings were combined and reversed-scored to create a negativity measure (α = .64). The three morality questions were combined into a wrongness index (α = .92). Finally, participants’ ratings of negative items from the PANAS were combined into a single negative affect scale (α = .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, β = .19, t(126) = 2.39, p = .02; it was not predicted by politics (β = .04), self-reported negative affect (β = -.04), story content (β = .04), or general negativity as assessed through attractiveness ratings (β = -.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 (\( rs < .18 \), \( ps > .30 \)). 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. 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. Finally, all participants completed a demographics survey as in previous studies.

Results and Discussion

The child ratings were averaged by scenario type (all \( a_s > .80 \)). 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 (\( M = -.11, SD = .55 \)), \( ps < .001 \). 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, \( r(99) = 2.73, p = .008 \). 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 \( rs < .13, ps > .21 \)).

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

\[ ^7 \text{Nor can this account for the link between immorality and harm for ostensibly harmless chicken masturbation or dog eating.} \]

\[ ^8 \text{An analysis of covariance is not appropriate here because each measure (suffering and boredom) has its own baseline.} \]
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-
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. 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 epistemic 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%). 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 consilience 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,
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—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


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.