THE UNIFYING MORAL DYAD: LIBERALS AND CONSERVATIVES SHARE THE SAME
HARM-BASED MORAL TEMPLATE

Chelsea Schein

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial
fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology
(Social Psychology) in the College of Arts and Sciences.

Chapel Hill
2016

Approved by:

Kurt Gray

Keith Payne

Tim Ryan

ABSTRACT

Chelsea Schein: The Unifying Moral Dyad: Liberals and Conservatives Share the Same Harm-
Based Moral Template
(Under the direction of Kurt Gray)

Do moral disagreements regarding specific issues (e.g., patriotism, chastity) reflect deep cognitive differences (i.e., distinct cognitive mechanisms) between liberals and conservatives? Dyadic morality suggests that the answer is "no." Despite moral diversity, we reveal that moral cognition—in both liberals and conservatives—is rooted in a harm-based template. A dyadic template suggests that harm should be central within moral cognition, an idea tested—and confirmed—through six specific hypotheses. Studies suggest that moral judgment occurs via dyadic comparison, in which counter-normative acts are compared with a prototype of harm. Dyadic comparison explains why harm is the most accessible and important of moral content, why harm organizes—and overlaps with—diverse moral content, and why harm best translates across moral content. Dyadic morality suggests that various moral content (e.g., loyalty, purity) are varieties of perceived harm and that past research has substantially exaggerated moral differences between liberals and conservatives.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The existence of moral disagreement across the political spectrum is uncontroversial. One need only open a news-paper to see that liberals and conservatives are divided on many issues, including abortion, capital punishment, gay rights, women's rights, gun ownership, environmentalism, euthanasia, and the justifiability of war. What is controversial is whether this disagreement reflects deep differences in moral cognition. Do liberals and conservatives have fundamentally different moral minds? One popular theory of moral cognition argues that liberals and conservatives rely on different sets of moral mechanisms (or foundations; Haidt, 2012).

In contrast, we suggest that liberals and conservatives fundamentally have the same moral mind. Rather than distinct and differentially activated mechanisms, we suggest that moral judgment involves a common template grounded in perceived harm (the moral dyad; Gray, Waytz, & Young, 2012). This template is not only consistent with categorization in non-moral domains but also reconciles modern moral pluralism with historic harm-centric accounts, and provides hope for bridging political differences. In this article, we test six predictions of dyadic morality, which can be summarized as follows: Harm is central in moral cognition for both liberals and conservatives.

Descriptive Diversity

An important first step in science is collecting and cataloging diversity. Biology began with natural history, in which living organisms were collected from around the world and placed into taxonomies. The most famous biological taxonomy is Linnaean classification—proposed by Carl Linnaeus—which divides organisms into five different kingdoms based on their appearance.

The new renaissance of moral psychology also began with collecting and taxonomizing moral diversity based on descriptive appearance, akin to Linnaean classification. Anthropological accounts of morality in rural India were divided into three content areas of autonomy, community, and divinity (Shweder, Much, Mahapatra, & Park, 1997). One later account inspired by American political disagreement—Moral Foundations Theory (MFT)—taxonomizes morality into the five content areas of harm, fairness, authority, in-group, and purity (Haidt, 2012).

Beyond providing a moral taxonomy, MFT also suggests differences in morality across liberals and conservatives, with only conservatives being concerned with authority, in-group, and purity. The idea of differences between liberals and conservatives is not new, as decades of research reveal that conservatives are more tolerant of inequality, are more religious, and believe more in a just world (for review, see Jost, Glaser, Kruglanski, & Sulloway, 2003). In particular, classic research on right-wing authoritarianism (RWA) reveals that conservatives are more submissive to authority, more likely to use aggression to protect the in-group, and more conventional in terms of sexuality (Altemeyer, 1988).

Given that RWA demonstrates the sensitivity of conservatives to authority, in-group, and sexual/religious conventionalism, it is safe for MFT to suggest the same—especially given that MFT questionnaires about authority, in-group, and purity correlate up to r = .70 with the RWA questionnaire (Graham et al., 2011, Table 7, p. 377). Moreover, as RWA fully accounts for liberal-conservative differences revealed by MFT (Kugler, Jost, & Noorbaloochi, 2014), one may wonder about the novelty of MFT claims. However, MFT does make one bold, unique claim—that these political differences arise from deep differences in moral cognition.

Moral Modules

Inspired by theories of basic emotions, MFT suggests that harm, fairness, in-group, authority, and purity each represent a distinct functional moral mechanism or cognitive module

(Haidt, 2012). MFT defines cognitive modules as "little switches in the brains of all animals" that are "triggered" by specific moral "inputs" (Haidt, 2012, p. 123). These modules are suggested to be ultimately distinct from each other, involving fundamentally "distinct cognitive computations" (Young & Saxe, 2011, p. 203), such that violations of one content area (e.g., harm) are processed differently from those of another (e.g., purity).

This "distinct cognition" prediction is best explained by the MFT analogy of moral foundations as different "taste receptors" (e.g., purity as "saltiness"), such that each moral concern triggers only one specific receptor, which gives rise to a corresponding distinct moral experience. However, recent evidence casts doubt on claims of distinct cognition, as even harm and purity—often discussed as maximally distinct (Haidt, 2012)—are highly correlated (r = .87; Gray & Keeney, 2015). Moreover, the apparent cognitive differences between these two content areas stem from scenario sampling bias: MFT purity violations are weirder and less severe (e.g., necrophilia) than harm violations (e.g., murder), and it is these general differences that give the illusion of distinct cognition (Gray & Keeney, 2015).

MFT also posits that moral modules are uniquely linked to specific emotions, such that harm is specially linked to anger and purity to disgust. Recent evidence also casts doubt on this claim, as links between moral content and emotion can be explained with broader affective and conceptual considerations (Cameron, Lindquist, & Gray, in press; Cheng, Ottati, & Price, 2013). Studies that purport to find a unique effect of disgust on purity fail to control for other high arousal emotions, such as anger or fear, or use statistical procedures that ignore substantial shared variance between anger and disgust (Cameron et al., in press).

Despite the lack of evidence for cognitive distinctness, the idea of moral "foundations" is intuitively compelling because it aligns with psychological essentialism—providing a deep

mental/biological explanation for important political differences. However, descriptive differences between liberals and conservatives—whether in personality, music preferences, or the moralization of specific issues—need not reflect cognitive differences. Even incredible diversity can be underlain by a common process.

Common Cognition

If science begins with cataloging diversity, its next step is often developing theories that explain this diversity with a common mechanism. In biology, Linnaeus and his contemporaries believed that each species was distinct and immutable, uniquely created by God. However, Darwin discovered that this incredible diversity stemmed from the simple algorithm of evolution. Could moral diversity also be underlain by a simple common mechanism? Could liberals and conservatives—despite their political disagreements—ultimately have the same moral cognition?

Decades of research in cognitive psychology suggest that non-moral categorization decisions (is x a member of y category?) rely upon the process of template comparison. This template (or prototype) represents the most common, salient or important features across category instances (Murphy, 2004). Categorization decisions are made by comparing potential examples with this template, with closer matches being more robustly categorized as belonging to that category. For the category of "birds," the template includes the features of "small," "flying," and "seed-eating," which explains why sparrows are judged as more bird-like than penguins (Rosch, 1978). The same process occurs in social categorization, in which people are compared with cognitive templates called stereotypes (Smith & Zarate, 1990).

The principle of parsimony suggests that moral judgments (is x act a member of the category immorality?) should also be made via template comparison. Acts should be compared to a moral template—or prototype—that extracts the most common, salient, and important

4

elements across instances of immorality. Moral psychology suggests many potential candidates

for these elements, such as concerns about intention, causation, and outcome (Alicke, 2000;

Cushman, Young, & Hauser, 2006; Malle, 2006); norm and affect (Nichols, 2002); and mind

perception (Gray, Young, & Waytz, 2012). The concept of harm is related to many of these

elements, and we suggest it forms the basis of a cognitive moral template.

Dyadic Morality

Harm can manifest itself in different ways, but within moral contexts, it typically

involves the intentional action of one person causing suffering to a second person—a perpetrator

and a victim. More technically, harm involves the perception of two interacting minds, one mind

(an agent) intentionally causing suffering to another mind (a patient)—what we call the moral

dyad (Gray, Waytz, & Young, 2012; see also Mikhail, 2007).

The complementary roles of agent and patient stem from the two-dimensional nature of

mind perception (Bastian, Laham, Wilson, Haslam, & Koval, 2011; Gray, Jenkins, Heberlein, &

Wegner, 2011) and the general dyadic structure of language (Brown & Fish, 1983) and action

(Aristotle, BC350), in which agents act upon patients (Strickland, Fisher, & Knobe, 2012). The

psychological power of a harm-based template stems not only from the presence of intentional

harm in many canonical acts of immorality (e.g., murder, rape, assault, and abuse) but also from

the affective potency of suffering victims (Blair, 1995), the hypersensitivity of agency detection

(Barrett, 2004), the early develop-ment of empathy and harm-based concerns (Decety & Meyer,

2008; Govrin, 2014; Hamlin, Wynn, & Bloom, 2007), and the obvious evolutionary importance

of harm (DeScioli & Kurzban, 2013).

It is clear that harm plays a key role in morality, helping to separate counter-normative

acts into those that are immoral from those that are violations of mere social convention (Sousa,

Holbrook, & Piazza, 2009; Turiel, Killen, & Helwig, 1987). Dyadic morality provides a mechanism for the role of harm. We suggest that a norm violation "x" leads people to automatically ask "is x immoral?"—per-haps to the degree that x induces negative affect (Nichols, 2002)—which then activates the dyadic harm-based tem-plate. The more an act is inherently dyadic (i.e., harmful), the better the template matches, and the more robustly it is judged as immoral, explaining why murder is judged as more immoral than masturbation. Importantly, this process of dyadic comparison is intuitive and need not rely on effortful reason, like moral judgment in general (Haidt, 2001).

Although a dyadic template should be reliably present during moral judgments, we acknowledge the influence of other domain-general psychological factors, such as misinterpreting affective arousal (Wheatley & Haidt, 2005) or rote learning (i.e., "the bible says abortion is wrong"). However, a dyadic template suggests that such misinterpretation and rote learning are easier with more harmful actions. Just as it is easier to rote-learn that a jerboa (a little desert kangaroo) is a mammal than a platypus (which non-prototypically lays eggs and has a bill), it should also be easier to rote-learn that abortion is immoral than to rote-learn that dropping the Torah is immoral.

Importantly, we are not suggesting that moral cognition consists of only one moral module (i.e., a foundation) of harm. Dyadic morality, with its roots in psychological costructionism (Cameron et al., in press), denies the very existence of moral modules. This template is not an on-or-off "switch" but is instead a domain-general process that allows for gradations of harm. It is also activated no matter the con-tent of the norm violation—that is, even when an act initially seems harmless (Gray, Schein, & Ward, 2014). Because harm represents the

essence of immorality, it serves as a constant backdrop in moral cognition—one that exerts a powerful cognitive gravity (Schein & Gray, 2014).

The Pluralism of Perceived Harm

Modular theories such as MFT have long argued against such a common template because of the ostensible existence of harmless wrongs. For example, scenarios of consensual incest carefully constructed to be "objectively harmless" are still rated as immoral by participants (Haidt, 2001). However, we argue against the very idea of "objective" harm. Harm, like morality, is in the eye of the beholder. In fact, both harm and morality are rooted in the ambiguous perceptions of other minds. Judgments of immorality require seeing a mind capable of doing evil, and judgments of harm require seeing a mind capable of suffering (i.e., an agent and a patient; Gray & Schein, 2012).

The subjective nature of harm means that bizarre "harm-less" scenarios concocted by liberal researchers (e.g., masturbating with a dead chicken) may not seem harmless to their more conservative participants. Indeed, many studies document the perception of harm in "harmless" cases of religious blasphemy, anti-patriotism, and aberrant sexuality (DeScioli, Gilbert, & Kurzban, 2012; Kahan, 2007; for a full treatment, see Gray et al., 2014).

Consider a case described by anthropologist Richard Shweder (2012): Oriya Hindu Brahmans believe it is extremely immoral for the eldest son to eat chicken immediately after his father's death. Westerners fail to see this action as wrong—or harmful—viewing it as a mere matter of religious protocol, whereas Hindus consider it the eldest son's duty to process the father's "death pollution" through a vegetarian diet. When the son eats chicken, he "places the father's spiritual transmigration in deep jeopardy" (Shweder, 2012, p. 96). By understanding the

perceived harm in these actions, even Western liberals can understand its perceived immorality.1 Who can deny the immorality of condemning your father to eternal suffering?

MFT interprets such perceived harm as mistaken, but dyadic morality sees these perceptions as legitimate. In the language of social anthropology, dyadic morality advocates for not only moral pluralism (accepting the legitimacy of different perceptions of morality) but also harm pluralism (accepting the legitimacy of different perceptions of harm). Harm pluralism suggests that different moral content such as purity and loyalty are (less prototypical) varieties of perceived harm. In contrast, MFT endorses harm monism, rejecting the legitimacy of harm in anything but direct physical or emotional suffering.

Indeed, the very act of separating harm into a specific modular "foundation" denies its perceived existence in moral issues such as treason or sexual impropriety. The harm monism of MFT discounts the harm that conservatives see in matters of religious and sexual propriety (Gray et al., 2014). We suggest that this harm monism stems from the liberal bias in social psychology (Inbar & Lammers, 2012), which also once long denied the legitimacy of moral pluralism. Echoing the cries of moral anthropologists, we suggest that understanding harm requires cultural sensitivity (Shweder, 2012); moral psychology should prioritize the harm pluralist perceptions of participants over the harm monist theories of researchers.

The Centrality of Harm for Liberals and Conservatives

The diversity of harm provides the possibility for a unify-ing moral template in both liberals and conservatives. Rather than distinct moral mechanisms for each kind of moral content, dyadic morality suggests that immoral acts—even those of "authority" or "purity"—will activate a prototype of harm. Of course, some acts are more harmful than others, and dyadic morality predicts that increased harm (i.e., better template matches) will result in more severe

8

judgments of immorality. Consistent with past research on RWA, we acknowledge political differences between liberals and conservative—and the possibility that these differences may translate to some differences in moral judgment. However, we predict that moral differences between liberals and conservatives have been greatly exaggerated by MFT (a prediction consistent with Frimer, Biesanz, Walker, & MacKinlay, 2013; Janoff-Bulman & Carnes, 2013; Skitka & Bauman, 2008; Skitka, Morgan, & Wisneski, in press).We suggest that liberals and conservatives share the same dyadic template, rather than categorically different moral minds. A harm-based moral template predicts that harm should be central in moral cognition across both moral diversity (i.e., many different moral acts) and political orientation (i.e., for both lib-erals and conservatives). Because centrality is a relatively broad concept, we operationalize it through six specific hypotheses:

Hypothesis 1 (H1: Accessibility): Harm is most cognitively accessible across moral diversity and political orientation (Study 1).

Hypothesis 2 (H2: Importance): Harm is most important across moral diversity and political orientation (Studies 2 and 3).

Hypothesis 3 (H3: Organization): Harm organizes judgments of immorality across moral diversity and political orientation (Study 4).

Hypothesis 4 (H4: Overlap): Harm overlaps substantially with other moral concerns across political orientation (Study 5).

Hypothesis 5 (H5: Translation): Harm is the best lingua franca for translating across moral diversity and political orientation (Study 6).

Hypothesis 6 (H6: Association): Harm is more implicitly associated with moral diversity than descriptively similar concerns, across political orientation (Study 7).

Accessibility (H1) Classic studies in cognitive social psychology reveal that prototypical examples are most cognitively accessible (Murphy, 2004). This accessibility explains why people who generate an example of a category (e.g., dog) generate an example that is prototypical (e.g., a Golden Retriever) rather than non-prototypical (e.g., a Xoloitzcuintli). Dyadic morality predicts that when both liberals and conservatives are asked to recall an example of "immorality," they will recall an action that is harmful, rather than "disloyal" or "impure" (Study 1).

Importance (H2) Research reveals that prototypical elements are most diagnostic for category judgments (e.g., the presence of wings is more diagnostic for whether something is a bird than the presence of eyes). The centrality of harm suggests that the pres-ence or absence of harm should be the most important criterion in moral judgments across moral diversity (Studies 2 and 3). In other words, relatively harmful acts should be most immoral, and relatively harmless acts should be least immoral.

We should note that maximal accessibility and importance of harm is technically consistent with distinct moral modules, such that harm might merely be the most important of multiple modules. However, MFT discusses moral concerns as if they are equally important and explicitly claims that "conservatives endorse all five foundations more or less equally" (Haidt, 2012, p. 187). Instead, we predict that both liberals and conservatives will view harm as the most important moral content, consistent with some recent research (Frimer et al., 2013). In contrast to statements of MFT that liberals have a deficient "two-foundation morality" (Haidt, 2012, p. 159), we also expect to find overall similarity in the endorsement of loyalty, authority, and purity between liberals and conservatives when these concerns are detached from specific issues (e.g., patriotism, gay marriage).

Organization (H3) One hypothesis that would distinguish "harm as common template" from "harm as most important module" is whether harm can meaningfully organize moral diversity. Distinct moral modules suggest a process of categorization (i.e., moral judgment) in which purity violations (e.g., bestiality) are evaluated by their impurity but not their harm. This separation of moral concerns means that perceptions of harm should not be able to meaningfully predict the immorality of non-harm transgressions.

Consider again the modular analogy of morality as distinct tastes (Haidt, 2012). If you had a set of foods that were purely sweet (and contained no salt), then it should not be possible to meaningfully organize these foods based on saltiness (beyond saying that they contained none). Said another way, the question "What is saltier: brown sugar or white sugar?" is meaningless, because they both entirely lack salt. Likewise, distinct moral modules suggest that the question "What is more harmful: bestiality or pornography?" is similarly meaningless because both are "purity" violations.

Conversely, harm pluralism and dyadic morality suggest that judgments of any potential moral infraction involve some perception of harm and the activation of a dyadic template. The process of dyadic comparison (i.e., template matching) suggests that even infractions of fairness, loyalty, authority, and purity are automatically organized on a gradient of harm. We test the organizational power of harm in both liberals and conservatives with an implicit judgment task, using speeded judgments of immorality and harm across moral diversity (Study 4).

Overlap (H4) Harm may meaningfully organize moral diversity, but one could argue that there may be many such potentially meaningful dimensions—perhaps one for each moral concern. Already, such an argument is inconsistent with strong modular accounts because dimensions represent fuzzy gradients rather than encapsulated mechanisms. Nevertheless,

suppose a weaker modular account that acknowledges the possibility of many dimensions, with harm being the most important. The question then becomes whether these dimensions are distinct from harm (and from each other). A key assumption of distinct moral modules is obviously distinctness, as such moral infractions are thought to activate one moral concern and not others (for a broader discussion, see Cameron et al., in press).

To test this claim of distinctness, we use scenarios specifically designed by MFT to activate one—and only one—type of moral content. Consistent with dyadic morality, we expect that distinctness will be lacking. Such conceptual overlap has already been observed in past research, as moral judgments between "foundations" are so highly correlated that many methodologists would consider them collinear (e.g., r = .87; Gray & Keeney, 2015). This overlap likely explains why factor analyses of moral foundation items only yield five factors when using confirmatory factor analyses with fit indices biased toward significance by large sample sizes (N = 34,476 in Graham et al., 2011).

If the overlap between moral content is sufficiently high for both liberals and conservatives, then we may question the predictive utility of treating fairness, loyalty, authority, and purity as separate concepts. Importantly, if harm is the most important dimension (H2), then, consistent with dyadic morality, substantially overlapping moral content may be best predicted and explained by harm. In other words, different moral content may be best understood as varieties of perceived harm.

Translation (H5) Substantial overlap between moral concerns—and the maxi-mal importance of harm—allows for the possibility that all moral infractions may be translated through a common language of harm. Modular accounts have suggested that different moral concerns represent different moral languages and that each language is distinct and ultimately

12

untranslatable (or more technically, incommensurable; Haidt, 2012). This means that it should be difficult or impossible, for example, to compare the relative immorality of a purity violation with that of a loyalty violation. Conversely, a common dyadic template suggests the possibility that various moral languages can be translated through harm for both liberals and conservatives (Study 6). Whether a purity or loyalty violation is worse is best answered by a simple question—Which transgression is seen as more harmful?

Association (H6) Some moral transgressions cluster together descriptively. For example, people's judgments regarding patriotism violations (i.e., loyalty) are good predictors of their judgments regarding sexual conduct (i.e., purity), because both are elements of RWA (Altemeyer, 1988). However, if diverse moral judgments are translated best through harm, it suggests that harm should be—cognitively speaking—more closely linked to descriptively dissimilar moral concerns (e.g., purity) than other descriptively similar concerns (e.g., loyalty; Study 7). We test implicit cognitive associations between moral content for both liberals and conservatives with the implicit association test (IAT; Greenwald, Nosek, & Banaji, 2003).

CHAPTER 2: Study 1: Recalling Immorality

Study 1 tests whether harm is most cognitively accessible for both liberals and conservatives (H1: Accessibility) using spontaneous recall. If a dyadic template is central to moral cognition, then dyadic violations (those involving an obvious agent and patient, such as murder, rape, and theft) should be spontaneously recalled with greater frequency than ostensibly harmless "purity" violations (e.g., sexual deviance).

Method

Participants

One hundred three participants completed the study on Amazon Mechanical Turk (mTurk). Twenty-two participants failed the instructional manipulation check, and 2 participants failed to follow instructions (instead listing random words), leaving 79 participants (51% male, Mage = 35, 56% liberal).

Procedure

Participants were asked to "list an act that is morally wrong" and to "write down whatever comes to your mind first." No other instructions were given to ensure that responses were as natural as possible. After the act was listed, it was looped back to participants, who then selected which one of five adjectives (harmful, unfair, disloyal, disobedient, and gross) it best represented. These adjectives were taken directly from moral foundations research (Graham, Haidt, & Nosek, 2009).2 To avoid influencing participants in advance of listing their immoral

act, they did not see these category labels until after they listed it. Then, participants completed demographics questions (politics ranging from 1 = strongly liberal to 7 = strongly conservative, gender, age, country) and the instructional manipulation check. Consistent with past work (Haidt, 2012), we define liberals here and else-where as those who respond 1 through 3 on the political scale and conservatives as those who respond 4 through 7.

Results

Words Recalled

See Figure 1 for a word map of recalled violations. More than 90% of recalled violations are dyadic in nature and display clear interpersonal harm (murder 44%, stealing 17%, adultery 11%, abuse 7%, cheating 5%, rape 4%, lying 2%).

Content Rating

Examining participant labels revealed that 68% of participants categorized their first act recalled as harmful, 9% labeled it as unfair, 14% labeled it as disloyal, 8% labeled it as disobedient, and 1% labeled it as gross. There was some diver-sity on how people categorized that same act (14 people recalled stealing, and of those, 5 labeled it as harmful, 4 unfair, 1 dis-loyal, and 4 disobedient). This diversity suggests overlap between moral content, arguing against strong claims of distinctness. A chi-square revealed that participants labeled their acts as harmful more than any other content area, $\chi2(1, N = 79) = 12.96$, $p < .001$, and this tendency did not differ by political affiliation, $\chi2(1, N = 79) = 0.59$, $p = .44$ (see Figure 2). As predicted by a dyadic template, canonically harmful violations were the most accessible for both liberals and conservatives.

CHAPTER 3: Study 2: Imagine Yourself As an Anthropologist

In this study, we test the importance of different moral concerns across liberals and conservatives by presenting participants with acts described as harmful, unfair, disloyal, disobedient, and impure. Using these content descriptions directly taps these moral concerns (e.g., purity) independently of political dis-agreement on specific issues (e.g., gay rights). Removed from specific issues, we expected overall similarity in moral judgments across liberals and conservatives. Most importantly, a dyadic template predicts that harm should be the most important predictor of moral wrongness because it is most representative of the category "immorality" (H2: Importance).

Method

Participants

One hundred eleven participants completed the study through mTurk. Nine participants failed the attention check, leaving 102 participants (39% female, Mage = 36, 51% lib-eral, all U.S. residents).

Procedure

Participants were asked to imagine themselves as an anthropologist studying a foreign tribe. One day, they overhear that a tribesmen has performed an act that is harmful, unfair, disloyal, disobedient, or impure. For example, the description for harm was "the tribesman performed a harmful action that caused others to suffer either emotionally or physically." These content descriptions were taken directly from moral foundations research (Graham et al., 2009). After each story, participants rated the immorality and wrongness of each act and whether the

tribesman should be punished, all on 5-point scales from Definitely Not/Not at All (1) to Definitely Yes/Extremely (5), which were collapsed into an immorality index (all αs > .76). Participants rated all five scenarios in random orders then completed demographics questions as in Study 1.

Results and Discussion

A 5 (within: Act Description) × 2 (between: Political Affiliation) mixed ANOVA revealed a significant main effect of act, $F(4, 400) = 72.74$, $p < .001$, $\eta 2 = .42$, such that both liberals and conservatives rated harmful acts as significantly more immoral than unfair, disloyal, disobedient, and impure acts, all $ps < .001$, with the exception of unfair in liberals, $p = .14$ (see Figure 3). Given that unfair violations are fundamentally dyadic (see Study 1), these data are consistent with a dyadic template. There was no main effect of politics, $F(1, 100) = 0.70$, $p = .41$, $\eta 2 = .007$, but there was a significant interaction between act and politics, $F(4, 400) = 3.98$, $p = .004$, $\eta 2 = .04$. This interaction was driven by a single difference: Consistent with past theorizing (Altemeyer, 1988; Graham et al., 2009), conservatives ($M = 3.19$, $SD = 0.93$, confidence interval [CI] = [2.91, 3.48]) rated standards of impurity and decency as relatively more important than liberals ($M = 2.71$, $SD = 1.08$, CI = [2.44, 2.99], $p = .02$). We should note that this difference was not replicated in the next study.

Consistent with a dyadic template, harm appeared to be the most important moral concern across moral diversity for both liberals and conservatives.3 In addition, the political differences revealed were much smaller than might be expected. Liberals do not only care about harm and fairness but also endorse disloyalty and disobedience at similar levels as conservatives (see Frimer et al., 2013; Janoff-Bulman & Carnes, 2013). In this study, liberals also endorsed purity concern at 80% of the extent of conservatives. In other words, despite claims that liberals and

17

conservatives have fundamentally different moral considerations, we observed overwhelming

similarity when these considerations were assessed directly rather than via specific political

issues (e.g., patriotism, chastity).

CHAPTER 4: Study 3: X but Not Y Judgments

The previous study revealed the importance of harm in moral cognition. However, as past research finds that people see harm in even ostensibly "harmless" purity and loyalty viola-tions (Gray et al., 2014), we may not have best isolated perceptions of harm. Therefore, in this study, participants read about the acts in Study 2 described as possessing one moral content but not other moral content (e.g., disloyal but not harmful). Consistent with Study 2, we predicted to observe the overall importance of harm in moral judgment (H2: Importance) and find few descriptive differences between liberals and conservatives.

Method

Participants

One hundred eleven participants completed the study through mTurk. Of the initial 111 participants, 6 people failed the attention check, leaving 105 participants (59% male, Mage = 33, 58% liberal).

Procedure

Participants rated 20 acts which were described as being X but not Y, with X and Y referring to MFT definitions of harm, unfairness, disloyalty, disobedience, and impurity. For example, the harm but not fairness scenario was "the tribesman performed a harmful action that caused others to suffer either emotionally or physically. However, the tribes-man did not act unfairly and people were not denied their rights." Participants assessed immorality as in Study 2, and these three items were combined into an immorality index ($\alpha = .82$).

Results and Discussion

To directly compare the importance of each kind of moral content on moral judgments, we took the difference between two complementary stories. For example, immorality ratings for "harm but not disloyal" minus those from "disloyal but not harm" yield the unique power of harm versus disloyalty to predict immorality.4 To analyze these "asymmetry scores," a one-sample t test was used with 0 as the test value (i.e., 0 connotes no asymmetry such that each type of moral content is equally important for moral judgments). As predicted, harm was a better predictor of immorality than all other moral content, ts(104) > 3.19, ps < .003. These t-values were converted into odds ratios to reveal exactly how much more important harm is compared with other content (see Figure 4). For example, harm is 5 times more important than impurity and 8 times more important than disloyalty in predicting judgments of immorality.

Contrasting liberals and conservatives revealed only 1 difference among asymmetry scores (out of a possible 10): Conservatives placed more emphasis on disloyalty versus unfairness than liberals, t(103) = 2.69, p = .008, a relative difference consistent with past research (Graham et al., 2009). However, other similarly suggested moral differences across liberals and conservatives in disobedience and purity were not replicated (nor were the differences in purity found in Study 2). As in Study 2, we found overwhelming similarity between liberals and conservatives in their endorsement of descriptive concerns. More importantly, this study revealed that—to the extent that moral content can be separated5—harm is by far the more important predictor of immorality for both liberals and conservatives, consistent with a common cognitive dyadic template.

CHAPTER 5: Study 4: Reaction Times (RTs) to Immorality

The previous studies suggest that harm is both maximally accessible and important across political orientation and moral diversity. In this study, we assess whether harm can structure moral diversity by predicting immorality ratings across various acts (H3: Organization). Demonstrating such organization with even ostensibly "harmless" acts would suggest the operation of dyadic comparison (i.e., template matching) across moral diversity, just as a dyadic template predicts.

As dyadic comparison is hypothesized to be intuitive and automatic, we used an implicit measure—RT—to assess both perceptions of harm and immorality. For both liberals and conservatives, we predicted that RTs (i.e., categorization speed) of harm judgments would predict RTs of immorality judgments, even for ostensibly harmless actions (e.g., por-nography), and even when controlling for general negativity and general RT.

Method

Participants.

Ninety-five participants completed the study through mTurk. Four participants were excluded from analysis for failing to complete all trials, and 4 additional participants were excluded for inaccurately categorizing neutral words (e.g., wallpaper) as immoral, leaving 87 participants (52% female, Mage = 36, 62% liberal).

Procedure.

Participants rated 40 words—20 moral terms (e.g., murder, pornography), 10 negative control words (e.g., feces, germs), and 10 control words (e.g., furniture, jumping)—on three

criteria, including whether the act was immoral, whether the act was harmful, and whether the act was unpleasant.

RT studies are highly sensitive to factors such as word frequency and length, so "unpleasant" served as a control for both low-level word characteristics (i.e., familiarity and length) and general negativity. At the start of each trial, partici-pants saw 2 words on the top corners of their screens "Immoral/Not Immoral," "Harmful/Not Harmful," or "Unpleasant/Not Unpleasant." Participants then categorized each target word accordingly (presented in the center of the screen). To simplify the task, participants completed the trials in randomly pre-sented blocks (immoral, harm, and unpleasant) in which they categorized all 40 words according to a single criterion.

Results

RT data are particularly sensitive to noise, so RTs greater than 3,000 ms or under 300 ms were cut from analysis (1.3% of all trials) as these times reflect non-implicit reactions. Harmful RTs, unpleasant RTs, politics, and all interactions between these terms were entered as fixed factors in a random-intercept multi-level model predicting immorality RTs for the 20 immoral actions (see Figure 5). As predicted, harmful RT significantly predicted immorality RT, $\beta = .19$, $SE = 0.03$, $p < .001$, $CI = [0.13, 0.24]$, even when controlling for unpleasant RT, $\beta = .12$, $SE = 0.03$, $p < .001$, $CI = [0.06, 0.17]$. There was no main effect of politics, $\beta = -4.35$, $SE = 11.48$, $p = .71$, $CI = [-27.13, 18.48]$, and no significant interaction between politics and harm RT, $\beta = -.007$, $SE = 0.02$, $p = .68$, $CI = [-0.04, 0.03]$, or unpleasant RT, $\beta = .01$, $SE = 0.02$, $p = .45$, $CI = [-0.02, 0.05]$ (see Figure 6). The lack of difference across politics suggests that implicit moral cognition is the same across liberals and conservatives (consistent with past work, Gray et al., 2014; Wright & Baril, 2011), just as a common dyadic template predicts.

As a more stringent test of the organizational power of harm across moral diversity, we reran analyses without the "harmful" and "unfair" violations (new list: adultery, bestiality, betrayal, cannibalism, disrespect, incest, pornography, and prostitution). As before, harm RT significantly predicted the RT of immorality of these "harmless" violations, $\beta = .26$, $SE = 0.04$, $z = 5.80$, $p < .001$, $CI = [0.17, 0.35]$, when control-ling for politics and unpleasantness. In other words, harm provides an organizational framework for even "harmless" moral transgressions, consistent with the process of dyadic comparison.

CHAPTER 6: Study 5: Correlations Between Moral Content

Study 4 revealed that harm forms an organizing dimension across moral content, but there may be other such dimensions corresponding to other content (e.g., disloyalty). Notably, these dimensions are only important if they represent inde-pendent dimensions. Studies 2 and 3 revealed that harm is the most important consideration within morality, and so any overlap between harm and other content areas is therefore most meaningfully characterized by harm.

In this study, we test the distinctness of moral concerns by presenting participants with scenarios specifically designed by moral foundations researchers to exemplify a single moral con-cern (Graham et al., 2009). We then assessed the moral content of each story with words drawn directly from the moral foundations "dictionary" (Graham et al., 2009)—again chosen by MFT researchers to represent a single moral concern. Despite these targeted scenarios and descriptors, dyadic morality sug-gests that moral concerns will overlap substantially.

Method

Participants.

One hundred seven participants finished the study through mTurk. Fifteen participants failed the instructional manipulation check, and 4 failed the experimental manipulation check by claiming that moral violations were "friendly." These exclusions left 88 participants (Mage = 39, 51% male, Mage = 31, 57% liberal).

Procedure.

Participants rated 10 different acts, 2 drawn from each of five content areas: harm (kick dog, insult overweight person), fairness (steal from poor, hire only people of own race),

disloyalty (burn country flag, break from family), dis-obedience (cursing nation's heroes, throw tomato at politician), and impurity (eat dead dog, engage in dehumanizing performance art). Participants then rated how well 12 different adjectives describe the action. The adjectives consisted of two words from each content area: harm (harmful, cruel), unfair (unfair, prejudiced), disloyal (disloyal, treacherous), disobedient (disobedient, disrespectful), and impure (impure, perverted), as well as two positive terms (friendly, delightful). The scenarios and moral descriptors were taken directly from past work advocating for distinct moral modules (Graham et al., 2009). The two positive words were included as attention checks.

Results

The two descriptors from each content area were averaged, and correlations were calculated between all five moral con-tent areas. Fisher r-to-z transformations revealed that the cor-relations between liberals and conservatives did not differ significantly, all ps > .65, and so analyses included all par-ticipants together. Analyses revealed that all moral concerns are correlated with each other, p < .001. In fact, when all 10 morality adjectives are submitted to a reliability analysis, it yields α = .89, suggesting strikingly internal consistency among moral concerns (i. e., a single scale). See Tables 1 and 2 for a listing of the correlations between moral concerns.

Although all moral content correlations are significantly positive, the raw correlations may underestimate the true latent correlations between moral content. Raw correlation coefficients between two operationalizations of x and y are limited by the individual reliability of those operationalizations (see Table 1). As MFT moral judgment items possess only modest reliability (harm, α = .51; fairness, α = .40; loyalty, α = .46; authority, α = .60; purity, α = .75; Graham et al., 2011, Table 2, p. 372), we used the standard correction for attenuation to reveal

the latent correlation between moral content (Muchinsky, 1996).6 As can be seen in Table 2, distinctness appears to be lacking in these latent construct correlations.

CHAPTER 7: Study 6: Translating Across Moral Diversity

The maximal accessibility and importance of harm (Studies 1-3), combined with its organizing ability across moral content (Study 4), and the substantial overlap between harm and other moral concerns (Study 5), suggests that harm might be used as a lingua franca—or common currency (Bauman, Wisneski, & Skitka, 2012)—between different content (H5: Translation). When trying to compare across different moral concerns, harm may serve as a natural language for translation. Of course, other moral concerns (each of which involve perceived harm—see Study 5) may also allow for translation, but a dyadic template suggests that prototypically harm should be the best.

By analogy, hot peppers can vary along a number of taste dimensions (e.g., smokiness, sweetness, size, color), but they are more meaningfully compared with each other on the dimension of hotness (i.e., capsaicin). We tested the translational power of harm through comparative judgments of immorality and moral content between different scenarios. We predicted that differences in immorality between scenarios would be best predicted by differences in perceived harm—even for ostensibly "harmless" scenarios.

Method

<u>Participants</u>.

Eighty-seven participants completed the study through mTurk. Twenty-five participants failed the attention check, leaving 62 participants (58% female, Mage = 33, 56% liberal).

<u>Procedure</u>.

Participants compared the immorality of the same 10 MFT-generated acts as in Study 5, 2 drawn from each of five content areas (Graham et al., 2009). As a manipulation check, participants first read each story and selected the category that best fit the scenario (harm, unfair, disloyal, dis-obedient, and gross). Scenarios were then presented in pairs with participants rating which scenarios were more immoral on a 5-point scale from Definitely Action A (1) to Definitely Action B (5), with Neither (3) as a midpoint. One comparison asked, for example, whether it was more immoral to steal from the poor or eat your dead dog. Next, participants rated which action they thought was more harmful, unfair, dis-loyal, disobedient, and gross, using the same scale. Participants made 20 comparisons, 2 each between domains. The order of the stories within each comparison and the order of the comparisons were randomized.

Results

Data were analyzed with a multi-level model that nested comparisons between content areas within participants, with politics as a subject-level variable. A dyadic template predicts that relative judgments of immorality should be predicted best by relative judgments of harm. As predicted, harm had the highest correlation with judgments of immorality: $\beta = .44$, SE = 0.03, z = 17.25, p < .001; followed by unfair: $\beta = .29$, SE = 0.03, z = 10.45, p < .001; gross: $\beta = .12$, SE = 0.02, z = 5.36, p < .001; disobedient: $\beta = .06$, SE = 0.02, z = 2.83, p < .001; and disloyal: $\beta = .03$, SE = 0.02, z = 1.24, p = .21. Comparing βs with 95% confidence intervals revealed that harm (CI = [0.50, 0.38]) predicted immorality significantly more than all other content areas (see Figure 7).

For an even more stringent test of the translational power of harm, we removed all comparisons that included harmful scenarios (guarding against the possibility that its relative importance was driving both immorality and content judgments). The model revealed similar

results with harm best predicting immorality, $\beta = .44$, SE $= 0.03$, $z = 13.76$, $p < .001$, providing the clearest evidence for harm as a lingua franca because harm mediated between all scenarios that were not themselves harmful.

Politics did not significantly correspond with the immorality judgments when looking at all comparisons, or at only the "harmless" act comparisons, $\beta s < .02$, $ps > .44$. Although past research has found political differences in the ratings of some of these scenarios (Graham et al., 2009), relative judgments regarding which scenario was more immoral did not differ significantly by politics. There was, however, a significant interaction between politics and disloyalty, $\beta = .04$, SE $= 0.01$, $z = 2.28$, $p = .02$, suggesting that disloyal ratings were a better predictor for immorality ratings for conservatives than for liberals. Nevertheless, liberals were still sensitive to disloyalty, and there were no significant political differences observed for disobedience or impurity.

There are many potential metrics on which to compare moral acts, but a dyadic template suggests that (prototypic) harm should be the best. Consistent with this idea, the relative immorality of two acts was best predicted by harm across moral diversity, and for both liberals and conservatives. This further suggests that moral concerns are not incommensurate but can be understood through the language of harm.

CHAPTER 8: Study 7: Clustering IAT

Descriptively, some work suggests that moral diversity splits into two correlated clusters: Harm and fairness form an "individuating" cluster, whereas disloyalty, disobedience, and purity form a "binding" cluster (Graham et al., 2009). The high internal consistency of all moral concerns revealed in Study 5 suggests that these claims are substantially overstated.

Moreover, even descriptive clustering need not reflect psychological processes. A dyadic template suggests that diverse moral content should cognitively be linked to harm, even more than descriptively similar content. In other words, "impurity" should be more closely linked to "harm" than to "disloyalty," despite previous findings that endorsements of impurity and disloyalty are correlated across the political spectrum. Indeed, we suggest that these past correlations between politics and impurity/disloyalty stem from specific operationalizations of moral concerns that are biased toward conservatism, defining loyalty as patriotism and purity as sexual traditionalism.

As cognitive structure is best revealed by implicit tests, we used two IATs (Greenwald et al., 2003) to test associations between harm, loyalty, and impurity. All IATs involve two category distinctions (e.g., good/bad, black/white), and by looking at the relative pairing between these categories, inferences about implicit associations can be made. In the classic example, if participants are faster to categorize black faces when they are paired with bad (vs. good), it is evidence of an implicit association between the categories black and bad. In the case of morality, we contrast judgments of (a) immoral versus not-immoral and (b) harm-related content versus "non-harm" moral content (e.g., disloyalty). We predict that specific immoral acts (e.g.,

impurity-related acts) cluster better with harm than with non-harm moral content (e.g., disloyalty).

General moral pluralism is consistent with any pattern of judgment, because moral pluralism is a statement about descriptive differences and not cognitive mechanisms. However, typical interpretations of moral pluralism argue that these descriptive differences correspond to the cognitive structure of moral judgment (Haidt, 2012). Conversely, a dyadic template suggests that impure acts—to the extent that they seem immoral—should be more closely linked to harm than to disloyalty for both liberals and conservatives.

Study 7a: Impurity and Harm Versus Disloyalty

This IAT tested whether immoral items related to impurity were better linked to disloyalty (as predicted by descriptive analyses; Graham et al., 2011) or to harm (as predicted by dyadic morality).

Method

Participants. One hundred participants were recruited through mTurk. Twenty-one participants were excluded from analysis for incorrectly categorizing more than 20% of the trials, leaving 79 participants (56% female, Mage = 38, 47% liberal, all from the United States).

Procedure. Participants completed an IAT with Content: Harmful/Disloyal and Immoral: Immoral/Non-Moral as the categorization pairs. Harmful items included victim, harmful, and dangerous. The disloyal items were disloyal, unfaithful, and unpatriotic. To measure the implicit associations with purity violations, the immoral items were incest, bestiality, and prostitution, and the non-moral items were forget, procrastinate, and boring—all items that are negative in valence but non-moral to control for general negativity. Importantly, participants read descriptions about the immoral items that defined them as a priori harmless. For example, the

31

case of incest was modeled after the "harmless" consensual incest used by those advocating for moral modules (see Haidt, 2001, p. 814). After completing the IAT, participants filled out demographics information.

<u>Results</u>.

IATs were conducted with Millisecond by Inquisit, and D-scores were calculated automatically according to established guidelines (Greenwald et al., 2003). Positive D-scores indicate that impurity-related immorality was more associated with harm than with disloyalty, a fellow "binding foundation." Consistent with a dyadic template, the mean D-score was positive (D = .21, SD = 0.69), which a one-sample t test revealed was significantly different from zero, t(78) = 2.72, p < .01. The more people saw impurity acts as immoral (vs. non-moral), the more they appear to link them to harm over disloyalty. The correlation between the D-score and politics was not significant, r(77) = .03, p = .80, revealing similar effects for both liberals and conservatives.

Study 7b: Disloyalty and Harm Versus Impurity

The first IAT revealed that impurity-related immorality was better linked to harm than to disloyalty. However, because the IAT assesses pairs of concepts, the link between purity-related immorality and harm may possibly arise because the non-moral terms are linked to disloyalty. Although there is no a priori reason to believe that words such as "procrastinate" should be more associated with disloyalty than harm, these concerns are best addressed empirically. As a dyadic template suggests that disloyalty-related immorality should be better linked to harm than to impurity, we reran an IAT using these categories. If we again found a link between disloyalty-related immorality and harm, it would not be explained by links between impurity/disloyal and non-moral terms.

<u>Method</u>

Participants. One hundred participants were recruited through mTurk. Twenty-six participants were removed from analysis for low accuracy (failing at least 20%), leaving 74 participants (59% female, Mage = 37, 53% liberal).

Procedure. The procedures for this IAT were identical to the previous one, with two exceptions. First, the word gross replaced disloyal as the category opposing harmful, with the corresponding items disgusting, gross, and filthy. To mea-sure the implicit associations with loyalty-related moral violations, the immoral items were adultery, treason, and betrayal.

<u>Results.</u>

Data were analyzed as in the previous study. Positive D-scores indicate that in-group moral violations are more associated with harm than with purity, a fellow "binding foundation." The mean D-score was .21, SD = 0.63, which a one-sample t test revealed was significantly different from zero, $t(73) = 2.89$, $p < .01$. Thus, the more people saw disloyal acts as immoral (vs. non-moral), the more they linked them to harm over impurity. There was no significant correlation with politics. The correlation between the D-score and politics was not significant, $r(72) = -.07$, $p = .28$, revealing similar effects for both liberals and conservatives.

Discussion

Ratings of immorality of both purity and loyalty violations were more strongly associated with harm than with their fellow so-called "binding" content, consistent with a harm-based dyadic template. The lack of differences between liberals and conservatives is also consistent with a common dyadic template.

CHAPTER 9: General Discussion

Seven studies suggest that harm is central in moral cognition across moral diversity for both liberals and conservatives. Across various descriptive moral concerns (harm, fairness, disloyalty, disobedience, and impurity), harm is most accessible (H1: Accessible; Study 1) and most important (H2: Important; Studies 2 and 3). Harm can meaningfully organize moral judgments across moral diversity, predicting the wrongness of even "objectively harmless" acts (H3: Organization; Study 4). Although other moral concerns may also organize moral content, concerns of harm significantly overlap with other moral concerns—which also overlap with each other (H4: Overlap; Study 5). The maximal accessibility and importance of harm, along with its overlap and organizational ability, allows harm to be the best lingua franca when comparing between moral concerns (H5: Translation; Study 6). Implicit tests also reveal that harm is more closely associated with other moral concerns than even descriptively similar concerns (H6: Association; Study 7).

Together, these studies are more consistent with a common dyadic template than with a specific number of distinct moral mechanisms that are differentially expressed across liberals and conservatives. As various moral content substantially overlap with (the more important) harm, different moral concerns can be best understood as different forms of perceived harm. Importantly, these studies revealed overwhelming similarity in moral judgments across liberals and conservatives.

Caveats

Of course, one must be cautious in the claims made by any single set of studies. In these studies, we used a relatively narrow set of stimuli, and it is possible that more diverse stimuli will reveal a reduced centrality of harm. However, the stories, labels, and definitions used here provide the most challenging test of harm's centrality, as they were taken directly from past research that advocates for distinct moral modules (Graham et al., 2009).We must also acknowledge that we did not examine moral judgments across all cultures, but only Americans. Nevertheless, researchers advocating for distinct moral modules argue that liberals and conservatives represent different moral cultures (Graham et al., 2011). We did reveal some descriptive differences in the endorsement of loyalty (Studies 3 and 6) and purity (Study 2), consistent with this past work (Graham et al., 2011). However, these differences were inconsistent across studies and, even when found, represented only small relative differences (consistent with RWA; Altemeyer, 1988) and not categorically different moral judgments.

To put these findings in context, there were more than 25 analyses across seven studies that compared liberals and conservatives on disloyalty, disobedience, and purity. Of these, only 12% found any differences between liberals and conservatives. Compared with the overwhelming power of harm, these other differences were minor, with no results supporting past claims that liberals have only a "two-foundation" morality. Most importantly, any descriptive differences in explicit endorsement do not appear to reflect underlying cognitive differences, as implicit measures revealed no differences between liberals and conservatives.

We also acknowledge that underlying cognitive structures—whether a dyadic template or distinct moral modules—cannot be directly revealed. Nevertheless, we suggest that these studies lend evidence to dyadic morality, especially in light of other recent evidence. Studies on dyadic completion find that the dyadic template exerts top-down effects on moral judgments (Gray et

al., 2014), such that any acts judged to be immoral are automatically perceived as harmful—even if they are "objectively harmless."

Together, the processes of dyadic comparison (if x is harmful, x is wrong)—revealed here—and dyadic completion (if x is wrong, x is harmful) form a dyadic loop, a cognitive feedback loop that underlies moral judgment. This loop exerts cognitive gravity—that is, constraint satisfaction—that bends perceptions of harm and immorality to be mutually consistent (Gray & Schein, 2012; Schein & Gray, 2014; see also Ditto & Liu, 2011; see Figure 8). The dyadic loop suggests that harmful actions seem immoral, and immoral actions seem harmful—a feedback cycle that may help us understand both political polarization and moral acquisition. Small differences in the perceived harm of a norm violation can elevate moral judgment, which in turn can elevate perceptions of harm, and so on, until one person believes it is incredibly immoral (and harmful) to eat beef and another incredibly immoral (and harmful) to eat pork.

A Moral Compromise

A dyadic template represents a synthesis of two long conflicting positions: historical cognitive-developmental accounts suggesting that harm is the only legitimate moral concern (Turiel et al., 1987), and pluralist accounts suggesting moral diversity (Shweder et al., 1997). A common cognitive template within moral diversity affirms the legitimacy of descriptive moral diversity but simultaneously highlights the importance of harm within cognition. Nevertheless, this compromise may be unsatisfying for those deeply committed to each theory. Cognitive developmentalists committed to moral reasoning (Kohlberg, 1981) may chafe at the idea that moral judgment involves the automatic and intuitive process of dyadic comparison. Likewise, moral pluralists may chafe at the idea that diverse moral judgments activate a common cognitive template.

Beyond understanding the mechanism of cognition, these data also have implications for taxonomizing and describing moral diversity. Moral pluralists often discuss moral concerns as if they are distinct and equal, and one popular article suggests that harm, fairness, disloyalty, disobedience, and impurity are separate and equal (Graham et al., 2011). The current data suggest two ways this idea could be modified. First, these concerns—represented by circles—should not be equal sizes because different concerns vary in importance (Studies 2 and 3). Second, the circles should not be distinct but overlap with each other (Study 5). The relative size (area) of the circles can be calculated from the odds ratio from Study 3, and the overlap can be calculated from squaring ($R2$) the latent correlation (i.e., adjusted for attenuation/measurement error) between content from Study 5. As can be seen in Figure 9, these studies propose a very different descriptive moral landscape than suggested by past research.

Given dyadic morality's emphasis on perceived harm, some may wonder whether it sacrifices pluralism for parsimony, overlooking cultural and political differences. We suggest that dyadic morality is not only parsimonious but is actually more consistent with moral pluralism. Rather than harm monist modular accounts, dyadic morality advocates for harm pluralism and acknowledges the richness of perceived harm, viewing various moral content as varieties of perceived harm: Harm can stem from direct physical injury and emo-tional harm, to damaging society and your immoral soul.

Dyadic morality suggests that moral disagreement typically occurs through different informational assumptions about who or what is vulnerable to harm (Turiel et al., 1987). Atheists deny the existence of souls and so scoff at the idea of religious sin, whereas believers see the soul as legitimately vulnerable and so guard against disobeying God. Those who see the poor as incapable of suffering will judge redistribution of wealth to be immoral, and those who see

fetuses as babies rather than mere cells will judge abortion to be immoral. If conservatives do have a wider moral domain (a premise argued against by the current data), this may stem simply from the fact that conservatives see relatively more threat in the world (i.e., potential harm; Jost et al., 2003).Dyadic morality also acknowledges that liberals and conservatives legitimately differ on specific issues, but we suggest that modular accounts have misrepresented these differences. Research finds that MFT has used a biased sample of stimuli when assessing political variability in morality (Gray & Keeney, 2015), asking questions about rolling in urine and bizarre plastic surgery, but not taxation, gun control, euthanasia, capital punishment, or environmentalism.

How can a theory of moral disagreement ignore the most contentious and representative cases of disagreement? By using only issues outlined by RWA for loyalty, disobedience, and impurity (see Kugler et al., 2014), it is unsurprising that MFT studies revealed that conservatives care more about these ideas. Conservatives may care more about "purity" when it is defined as religiosity and traditionalism but likely not when it is defined as environmentalism or organic foods. Likewise, liberals may seem more individualistic regarding patriotism but likely not when concerning taxation or gun rights. Descriptive taxonomies are undoubtedly useful, but only when they accurately catalog all of moral diversity, rather than a small subset of specific issues known a priori to show political differences (Haidt, 2012).

Future Directions

Any new theory must provide directions for future research, and we suggest seven.

First, how does a template of harm translate into different judgments of right and wrong across cultures? The same question was asked of evolution: How can one com-mon process provide the richness of biological diversity? With evolution, the key is determining the selective

pressures of specific environments; with morality, the key may be determining what varieties of harm cultures face (or perceive they face). Cultures with high prevalence of sexual diseases likely moralize sexual chastity (van Leeuwen, Park, Koenig, & Graham, 2012), cultures with many food-borne illnesses likely develop food-related taboos, and cultures constantly at war likely moralize group cohesion. In addition to anthropological studies, lab studies could examine links between perceived threat and moralization. As discussed above, we suggest that the dyadic loop could spark a feedback cycle that amplifies small (and potentially random) initial cultural differences in perceived harm/immorality.

Second, what role does perceived harm play in acquiring morality? As adults, it may seem that some things seem wrong simply because of entrenched childhood learning, but does childhood learning involve perceptions of harm? Dyadic morality predicts that children most easily moralize obviously harmful acts and that parents' use of harm-based language will most successfully moralize other cultural conventions.

Third, what role does perceived harm play in judgments of moral character? Recent work suggests that morality often involves judgments of people rather than acts (Goodwin, 2015; Uhlmann, Pizarro, & Diermeier, 2015). Dyadic morality suggests that judgments of character ultimately revolve around potential harm. Although some studies find that "impure" acts (e.g., chicken masturbation) reflect more poorly upon character than "harmful" acts (e.g., theft; Uhlmann & Zhu, 2013), these impure acts are much weirder than the harmful acts (Gray & Keeney, 2015). We suggest that this weirdness leads others to view people as unpredictable and dangerous (Gray & Keeney, 2015).

Fourth, what is the role of motivational processes in dyadic morality? Dyadic comparison and completion are often dis-cussed as inevitable consequences of a dyadic template, but

motivation can likely accelerate or dampen these processes. Those who want to find a person blameworthy will likely see their actions as more harmful than someone who wants to for-give them.

Fifth, what is the best taxonomy of perceived harm? Different cultures clearly differ on what they see as vulnerable to damage (Shweder et al., 1997). What is the best system to represent these differences? One possibility is to focus on the identity of the recipient of harm, as Janoff-Bulman and Carnes (2013) do by separating harm to the future self (e.g., drug use), harm to a specific other (e.g., murder), and harm to the group (e.g., cheating on taxes). Alternatively, one may focus on the method of harm, such as physical, mental, or spiritual. Although MFT appears not to describe the cognitive processes underlying moral judgments, a reformulation of this theory may represent a useful taxonomy of perceived harm.

Sixth, how does the dyadic template translate into good-ness? Morality is not only about harm but also helping (i.e., alleviating the suffering of victims)—Is there a dyadic help-based template for virtue?

Seventh—and perhaps most importantly—how can a dyadic template help solve political conflict? If people all speak the same moral "language" of harm (Mikhail, 2007), we should be able to translate moral issues across partisan divides. More specifically, by granting legitimacy to our political opponent's perceptions of harm, we may be more accepting of their views and more willing to engage in discussion.

Conclusion

The eminent anthropologist Richard Shweder is an advocate of both moral pluralism and the idea of "moral universalism without the uniformity" (Shweder, 2012, p. 88). He suggests that people across cultures would arrive at the same moral judgment if they had a "full specification

of circumstances and context" (Wiggins, 1990; cited in Shweder, 2012, p. 95). What exactly are these circumstances and context? Perhaps perceived harm.

The moral dyad allows for "universalism without the uniformity," because it provides a common structure for understanding moral cognition, while allowing diversity in the perceived methods and recipients of harm. In con-trast to other theories that divide up moral judgments—and the cultures who make them—into discrete bins, the moral dyad unifies morality. Despite ubiquitous moral disagreement, each of us seems to share a common cognitive template. Inside the moral minds of both liberals and conservatives beats the heart of harm.

Table 1

Raw Correlations Between Five Moral Content
Areas (Study 5)

Content Area	1	2	3	4	5
1. Harm	—				
2. Fairness	.59	—			
3. Disloyalty	.59	.83	—		
4. Disobedience	.62	.73	.75	—	
5. Impurity	.55	.44	.55	.49	—

Table 2

Latent Correlations Between Five Moral Content Areas,
Corrected for Attenuation (Study 5)

Content Area	1	2	3	4	5
1. Harm	—				
2. Fairness	1.0	—			
3. Disloyalty	1.0	1.0	—		
4. Disobedience	1.0	1.0	1.0	—	
5. Impurity	.89	.81	.93	.73	—

Figure 1. Free-recalled words from Study 1 across both liberals and conservatives. Size of word reflects frequency of recall.[1]

---
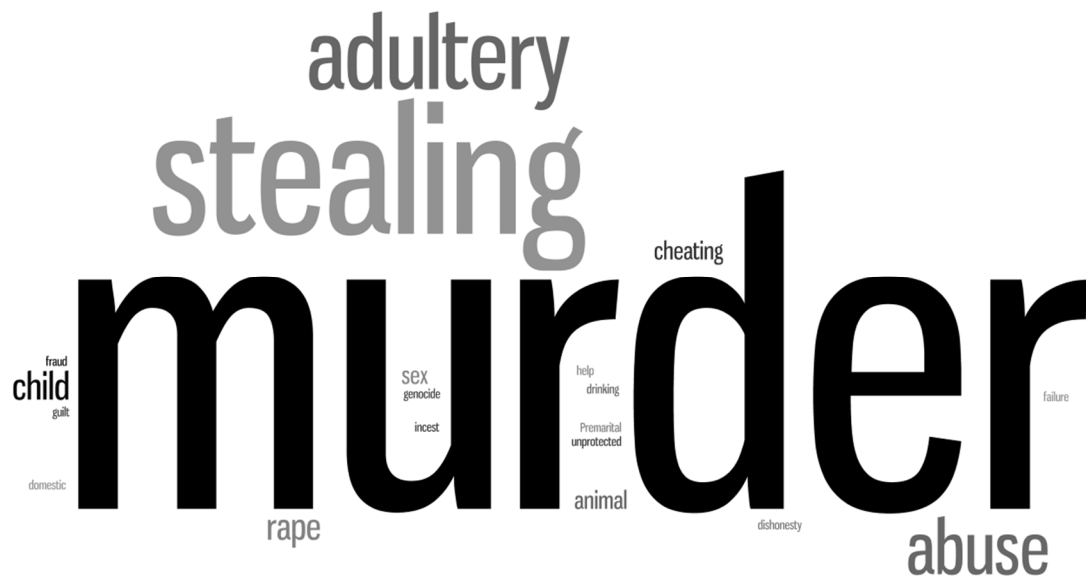
[1] Given their similarity, killing was combined with murder; theft with stealing; infidelity, cheating on spouse, and affair with adultery; punching/ kicking a child with child abuse

44

Figure 2. Participant categorization of free-recall act (Study 1).
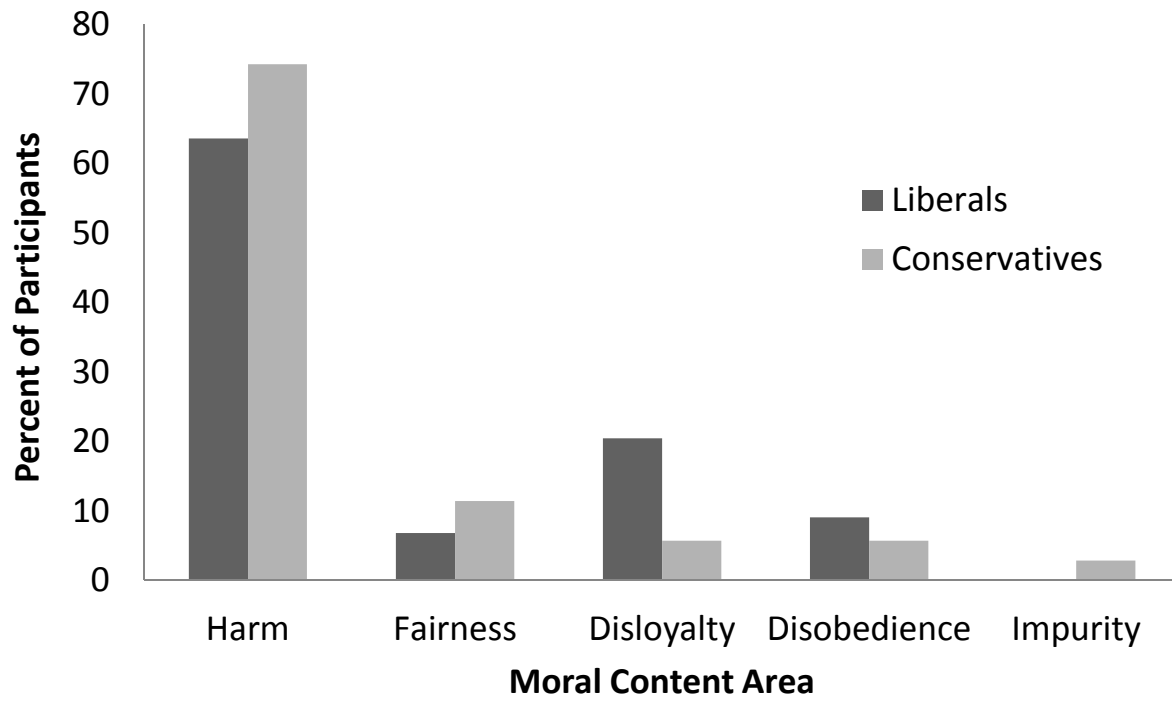
Figure 3. Perceived wrongness of various moral content for both liberals and conservatives
(Study 2). Error bars are +/-1 Std. Err.
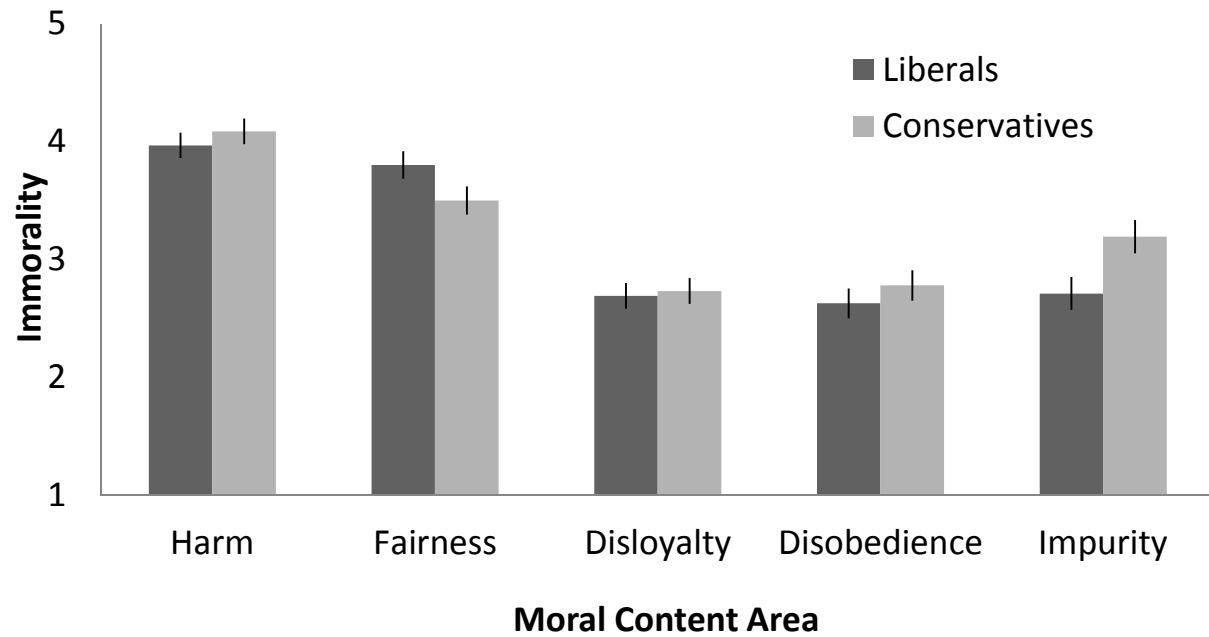
Figure 4. Comparing the importance of moral content area with each other via asymmetry scores (Study 3). Included are *t* values and *odds ratios*—which give the relative importance of moral concern (e.g., Harm is 8.32 times more important than disloyalty). Scores not listed are not significant (e.g., there is no difference in importance between disloyal and disobedience)
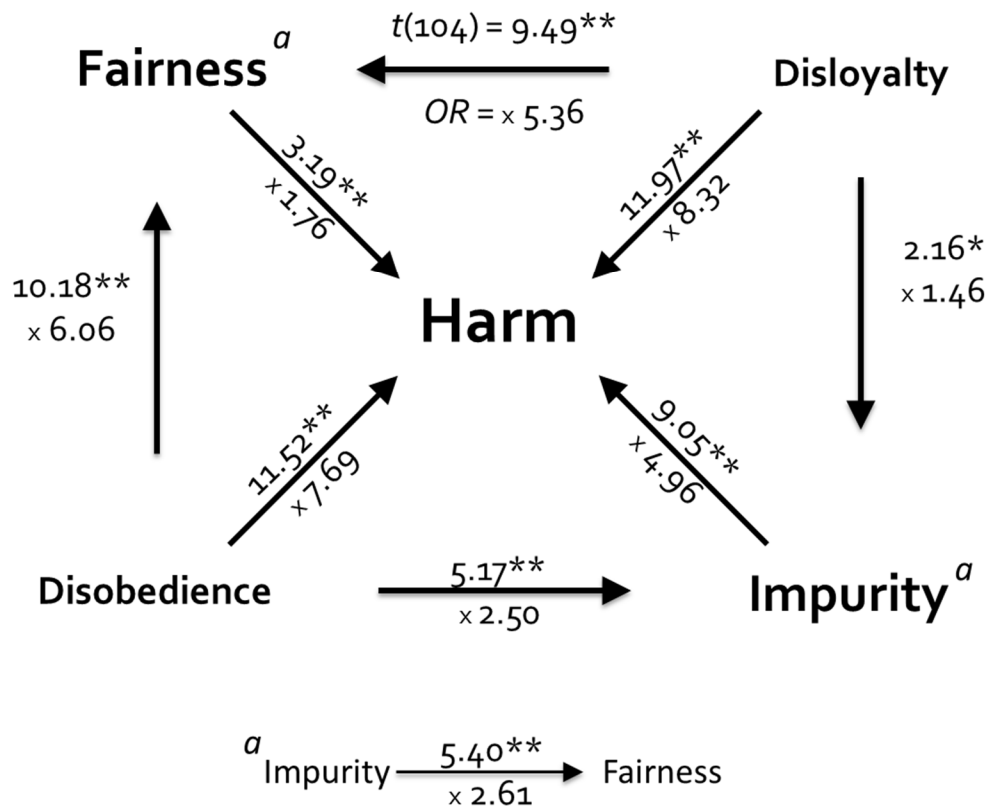
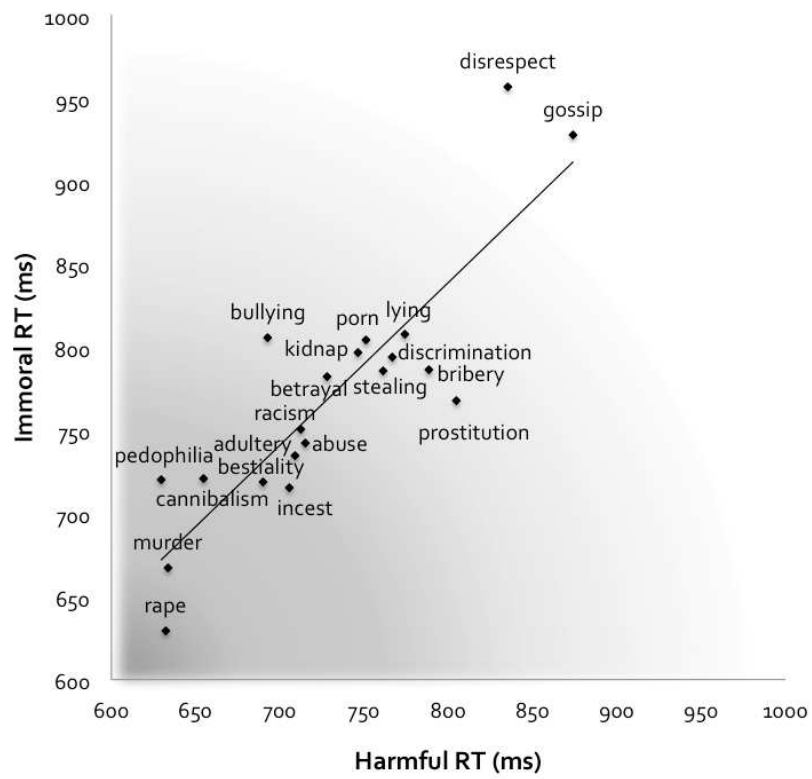Figure 5. Average time for categorizing an act as immoral and harmful by item (Study 4).

Figure 6. Average time for categorizing an act as immoral and harmful, split by politics (Study 4).
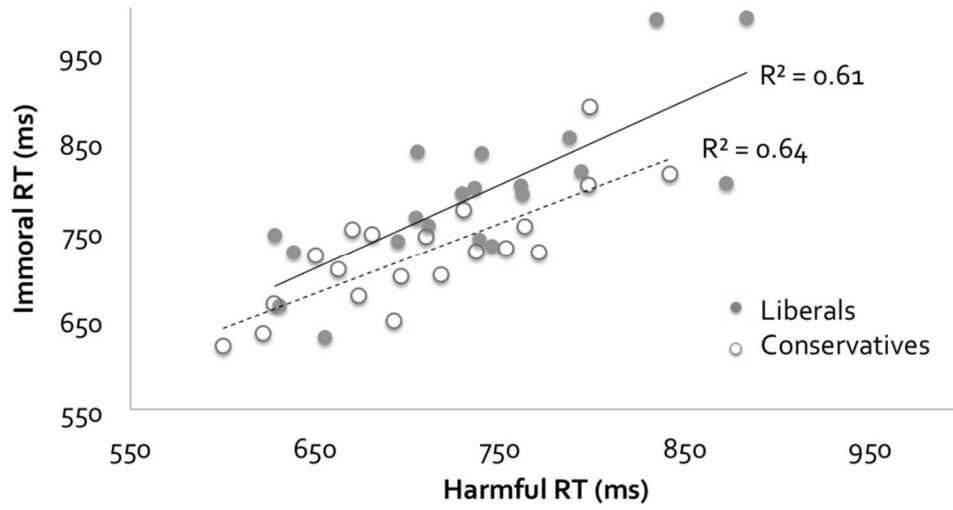
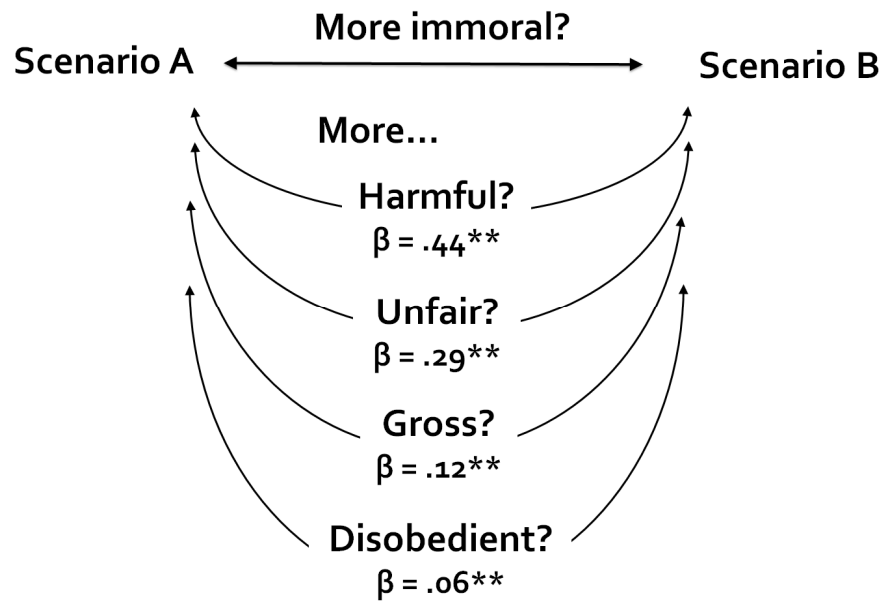Figure 7. The ability of various moral content to serve as a moral common currency (Study 6).

Figure 8. The Dyadic Loop. The bottom-up and top-down feedback cycle of the dyadic template.
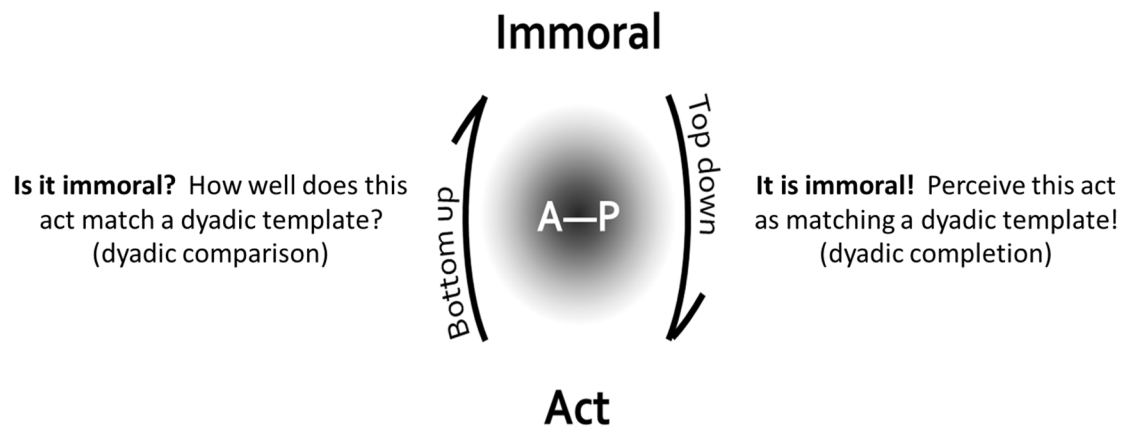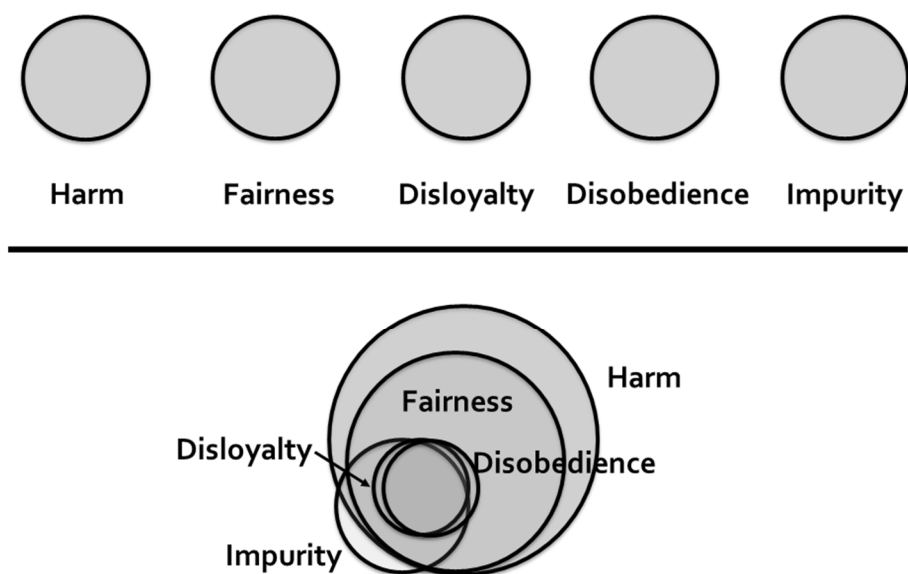
Figure 9. Descriptive moral content implied by moral foundations theory (top), contrasted with that revealed by Studies 3 and 5 (bottom).

REFERENCES

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*(4), 556–574.

Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism.* San Fransisco, CA: Jossey-Bass.

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

Asch, S. E. (1952). *Social Psychology* (1st ed.). Englewood Cliffs, N.J.: Prentice-Hall.

Barrett, J. L. (2004). *Why Would Anyone Believe in God?* Walnut Creek, CA: AltaMira Press.

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

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

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

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*(3), 237–273. http://doi.org/10.1016/0010-0277(83)90006-9

Cameron, C. D., Lindquist, K. A., & Gray, K. (in press). A constructionist review of worality and emotions: No evidence for specific correspondences between discrete emotions and moral concerns. *Personality and Social Psychology Review*.

Chapman, H. A., & Anderson, A. K. (2013). Things rank and gross in nature: a review and synthesis of moral disgust. *Psychological Bulletin*, *139*(2), 300.

Cheng, J. S., Ottati, V. C., & Price, E. D. (2013). The arousal model of moral condemnation. *Journal of Experimental Social Psychology*, *49*(6), 1012–1018. http://doi.org/10.1016/j.jesp.2013.06.006

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

Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, *20*(4), 1053.

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

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*(2), 477–496. http://doi.org/10.1037/a0029065

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.

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.

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, *24*(1), 38–44.

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

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*(5), 1029–1046. http://doi.org/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*(2), 366–385. http://doi.org/10.1037/a0021847

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

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

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

Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143*(4), 1600–1615. http://doi.org/10.1037/a0036149

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

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

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.

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

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

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. http://doi.org/10.1038/nature06288

Inbar, Y., & Lammers, J. (2012). Political diversity in social and personality psychology. *Perspectives on Psychological Science*, *7*(5), 496–503.

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape moral motives and group-based moralities. *Personality and Social Psychology Review*, *17*(3), 219–236. http://doi.org/10.1177/1088868313480274

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin;Psychological Bulletin*, *129*(3), 339–375. http://doi.org/10.1037/0033-2909.129.3.339

Kahan, D. M. (2007). The cognitively illiberal state. *Stanford Law Review*, *60*(1), 115–154.

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

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

Malle, B. F. (2006). Intentionality, morality, and their relationship in human judgment. *Journal of Cognition and Culture*, *6*(1-2), 87–112.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, *11*(4), 143–152. http://doi.org/10.1016/j.tics.2006.12.007

Muchinsky, P. M. (1996). The Correction for Attenuation. *Educational and Psychological Measurement*, *56*(1), 63–75. http://doi.org/10.1177/0013164496056001004

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

Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, *84*(2), 221–236.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.

Schein, C., & Gray, K. (2014). The prototype model of blame: Freeing moral cognition from linearity and little boxes. *Psychological Inquiry*, *25*, 236–240.

Shweder, R. A. (2012). Relativism and Universalism. In D. Fassin (Ed.), *A Companion to Moral Anthropology* (pp. 85–102). John Wiley & Sons, Ltd.

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

Skitka, L. J., & Bauman, C. W. (2008). Moral conviction and political engagement. *Political Psychology*, *29*(1), 29–54. http://doi.org/10.1111/j.1467-9221.2007.00611.x

Skitka, L. J., Morgan, G. S., & Wisneski, D. C. (in press). Political orientation and moral conviction: A conservative advantage or an equal opportunity motivator of political engagement? In *Social psychology and politics*. New York: Psychology Press.

Smith, E. R., & Zarate, M. A. (1990). Exemplar and prototype use in social categorization. *Social Cognition*, *8*(3), 243–262. http://doi.org/10.1521/soco.1990.8.3.243

Sousa, P., Holbrook, C., & Piazza, J. (2009). The morality of harm. *Cognition*, *113*(1), 80–92. http://doi.org/10.1016/j.cognition.2009.06.015

Strickland, B., Fisher, M., & Knobe, J. (2012). Moral structure falls out of general event structure. *Psychological Inquiry*, *23*(2), 198–205. http://doi.org/10.1080/1047840X.2012.668272

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. *The Emergence of Morality in Young Children*, 155–243.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*(1), 72–81. http://doi.org/10.1177/1745691614556679

Uhlmann, E. L., & Zhu, L. [Lei]. (2013). Acts, persons, and intuitions person-centered cues and gut reactions to harmless transgressions. *Social Psychological and Personality Science*. http://doi.org/10.1177/1948550613497238

Van Leeuwen, F., Park, J. H., Koenig, B. L., & Graham, J. (2012). Regional variation in pathogen prevalence predicts endorsement of group-focused moral concerns. *Evolution and Human Behavior*, *33*(5), 429–437. http://doi.org/10.1016/j.evolhumbehav.2011.12.005

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*(10), 780–784. http://doi.org/10.1111/j.1467-9280.2005.01614.x

Wiggins, D. (1990). Moral cognitivism, moral relativism and motivating moral beliefs. In *Proceedings of the Aristotelian Society* (pp. 61–85). JSTOR.

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. http://doi.org/10.1016/j.jesp.2011.03.014

Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, *120*(2), 202–214. http://doi.org/10.1016/j.cognition.2011.04.005