Harm Mediates the Disgust-Immorality Link

Chelsea Schein[1], Ryan Ritter[2], & Kurt Gray[1]

[1]University of North Carolina, Chapel Hill
[2]University of Illinois, Urbana-Champaign

Abstract

Many acts are disgusting, but only some of these acts are immoral. Dyadic morality predicts that disgusting acts should be judged as immoral to the extent that they seem harmful. Consistent with this prediction, three studies reveal that perceived harm mediates the link between feelings of disgust and moral condemnation—even for ostensibly harmless "purity" violations. In many cases, accounting for perceived harm completely eliminates the link between disgust and moral condemnation. Analyses also reveal the predictive power of anger and typicality/weirdness in moral judgments of disgusting acts. The mediation of disgust by harm holds across diverse acts including gay marriage, sex acts, and religious blasphemy. Revealing the endogenous presence and moral relevance of harm within disgusting-but-ostensibly-harmless acts argues against modular accounts of moral cognition such as moral foundations theory. Instead, these data support pluralistic conceptions of harm and constructionist accounts of morality and emotion. Implications for moral cognition and the concept of "purity" are discussed.

Key Words: Dyadic Morality, Ethics, Disgust, Purity, Values

Harm Mediates the Disgust-Immorality Link

Disgust has long been considered important in morality—both in public discourse and moral psychology. Bioethicist Leon Kass argued against stem cell research and cloning by appealing to the "wisdom of repugnance"(1997), and psychologist Paul Cameron argued against same-sex marriage by providing a disgusting account of gay sex involving "exchanging saliva, feces, semen and/or blood with dozens of different men each year" (Cameron, 2009 quoted in Nussbaum, 2010, p. 1). Reflecting this societal language of disgust, some scholars suggest a special psychological link between disgust and moral judgment (Horberg, Oveis, Keltner, & Cohen, 2009), but many disgusting acts are not immoral. Conventional heterosexual sex within marriage also involves an exchange of bodily fluids, possible contact with blood, and pungent odors, but is seldom judged as immoral. If feelings of disgust are not intrinsically or uniformly linked to moral judgment, then the question is what transforms "gross" into "wrong?" Perhaps harm. Drawing from the theories of dyadic morality (Gray, Young, & Waytz, 2012) and emotion construction (Cameron, Lindquist, & Gray, 2015), we suggest that perceptions of harm mediate the link between feelings of disgust and moral condemnation, even for ostensibly harmless acts.

Debates about disgust, immorality and emotion

Substantial research links moral condemnation to feelings of disgust—including individual differences (Crawford, Inbar, & Maloney, 2014; Inbar, Pizarro, Knobe, & Bloom, 2009) and experimental manipulations (Inbar, Pizarro, & Bloom, 2012; Terrizzi Jr, Shook, & Ventis, 2010)—but the nature of this association is debated (Pizarro, Inbar, & Helion, 2011). One debate involves the scope of disgust, with some advocating for links between disgust and all moral judgment (Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005), and others restricting the role of disgust to violations of bodily and/or spiritual "purity" (Haidt, 2012;

Horberg, Oveis, Keltner, & Cohen, 2009; P. S. Russell & Giner-Sorolla, 2013). Another debate involves the power of disgust, with some suggesting that disgust is sufficient for moralization—turning the nonmoral into immoral (Horberg et al., 2009)—and others suggesting that disgust merely amplifies pre-existing moral judgments (Pizarro et al., 2011).

Adding to these debates, a recent meta-analysis questions the very link between disgust and moral condemnation: across all published experiments, incidental manipulations of disgust have such a small impact upon moral condemnation as to be statistically non-significant after controlling for publication bias (Landy & Goodwin, 2015). We suggest that these debates are fueled by an intuitive but erroneous assumption: that disgust is a single thing.

Historically, moral psychology has been grounded in the theory of basic emotions (Graham et al., 2013; Rozin, Lowery, Imada, & Haidt, 1999), which argues for the existence of a number of discrete, domain-specific, culturally universal and ultimately coherent emotions (Ekman, 1992), each of which are linked to consistent and specific facial expressions, physiology and behavior (Ekman & Cordaro, 2011). Basic emotions theory suggests that one instance of disgust is much the same as another, and that—despite various "triggers"—disgust has a relatively unitary, uniform experience with consistent links to moral judgment (Graham et al., 2013). Although compatible with essentialist intuitions (Haslam, 1998), this theory has been challenged by recent research revealing little specificity or consistency in emotional experience and expressions across cultures (Gendron, Roberson, van der Vyver, & Barrett, 2014), in physiological activity (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000), brain activity (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), or behavior (Barrett, 2006)—all research consistent with theory of constructionism.

In contrast to basic emotions, constructionism suggest that emotions are emergent

combinations of the basic ingredients of affect (valence and arousal) and conceptual content

(Barrett, 2006; Lindquist, 2013; J. A. Russell, 2003). For example, high arousal negativity

transforms into fear through conceptual activation of threat, and into anger through the

conceptual activation of rights violations (Lindquist & Barrett, 2008). Constructionism denies

that emotions (e.g., fear) are unitary constructs caused by isomorphic mechanisms (i.e., a "fear"

module/mechanism/circuit). Instead, it suggests that emotion labels such as "fear" or "anger"

each represent a heterogeneous collection of experiences with corresponding heterogeneous links

to behavior and judgment (Barrett, 2013; Lindquist, 2013). Construction acknowledges that the

white-hot anger of being cut off in traffic involves substantially different experiences,

physiology, expressions and behaviors than the cold fuming of being insulted by your boss.

Consistent with constructionism—and recent work on disgust (Tybur, Lieberman, &

Griskevicius, 2009)—we suggest that the label of "disgust" also applies to a diverse collection of

experiences, such that people feel differently when contemplating hair in their food versus

masturbating with a corpse. This variability can help explain the variable association between

disgust and moral judgment—why only some instances of disgust are linked to moral

condemnation (Pizarro et al., 2011; Tybur, Merriman, Hooper, McDonald, & Navarrete, 2009).

The key question for moral psychology is then what distinguishes experiences of "moral

disgust" from "non-moral disgust?"   If disgust is constructed from more basic ingredients, then

which ingredient is reliably linked to immorality? One possibility is arousal (Cheng, Ottati, &

Price, 2013) or negative affect (Valdesolo & DeSteno, 2006), as both these elements are linked

to moral condemnation. However, most experiences of disgust seem to be arousing and

negative,[1] providing little discrimination among disgusting acts. As disgust is often linked to the

identification of threats (Crawford et al., 2014), another possibility is that disgust is linked to

immorality to the extent that it engenders perceptions of harm—an idea consistent with dyadic morality.

Dyadic morality

Dyadic morality suggests that moral judgment is constructed, just like emotion. Rather than combinations of affect and conceptual knowledge, dyadic morality combines two perceived minds: an intentional agent causing damage to a vulnerable patient (Gray, Young, et al., 2012). Together, this combination represents harm. Of course, as with disgust, there are many kinds of harm (e.g., stubbing your toe, car accidents), but we suggest that *morally relevant* harm is dyadic, involving intention + causality + suffering. We also suggest that this harm-based dyad forms a cognitive working model (or prototype) of morality (Schein & Gray, 2015).

More specifically, dyadic morality suggests that moral judgment (i.e., is *x* immoral?) is similar to non-moral categorization judgments (e.g., is *x* a bird?), involving the process of template comparison. Potentially immoral acts are compared to a dyadic template and are judged as immoral to the extent that they appear to "match" this template by involving intentional agents, suffering patients, and a causal connection between them (Gray & Schein, 2012; Mikhail, 2007). Dyadic morality can therefore explain why child abuse is seen as more immoral than double-parking, and why stealing from children is worse than tax fraud (Schein, Goranson, & Gray, 2015).

Importantly, the harm of dyadic morality is not the objective (Haidt, 2001), deliberative (Turiel, Killen, & Helwig, 1987), and monist harm (Haidt & Joseph, 2004) of historical moral psychology. Instead, it is subjective, intuitive and pluralist. This *perceived harm* varies across people and situations, tracking with different perceptions of immorality across cultures (Shweder, Much, Mahapatra, & Park, 1997). For example, Brahmins (but not Westerners) see

harm when funeral protocols are violated and Brahmins (but not Westerners) reliably judge these protocol violations as immoral (Shweder, 2012). Research finds that harm is automatically perceived in "objectively" harmless situations such as consensual incest (Gray, Schein, & Ward, 2014), and these automatic perceptions are so deeply intuitive that they defy explicit argumentation (Royzman, Kim, & Leeman, 2015). No matter the ostensible reasons experimenters concoct for incest being fun, safe and beneficial (Haidt, Bjorklund, & Murphy, 2000), it is hard to shake its perceived harmfulness—likely because these perceptions of harm serve an evolutionary function (Lieberman, Tooby, & Cosmides, 2003).

Just as with emotion construction, dyadic morality focuses more upon the *process* of moral cognition rather than its *content*. Accordingly, it acknowledges the practical utility of different taxonomies (Graham et al., 2013; Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011), and the experiential force of moral categories, such as loyalty (Haidt, 2012), equality (Brosnan & de Waal, 2003; Rai & Fiske, 2011), and social order (Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011). Murder clearly feels different from adultery, which feels different from embezzlement, but each can be understood through a dyadic template by varying the moral patient (victim, spouse, company) and the method of harm (physical, emotional, economic). Thus, despite an infinity of descriptively different moral concerns, each represents a manifestation of perceived harm (Gray, Waytz, & Young, 2012).

As an analogy, consider origami. Just as a single sheet of paper can be transformed into qualitatively different shapes, so too can a single template of harm be transformed into many qualitatively different concerns. "Impure" sexual violation and "unfair" economic violations clearly have different "shapes" (just like origami swans and foxes), but dyadic morality suggests that both these moral violations are transformations of the same underlying substrate—perceived

harm. This harm is not merely metaphorical. Instead, descriptively different concerns—even those that seem "harmless"—remain indelibly linked to perceived suffering, such as physical pain and emotional damage (DeScioli, Gilbert, & Kurzban, 2012; Gray et al., 2014).

Consistent with this idea, people seem to care about loyalty, equality, social order and other moral concerns because their violation seems to cause suffering—most often of children, who are characteristically vulnerable (Schein et al., 2015; Schein & Gray, in press). The suffering of children are referenced when arguing against gay rights (Bryant, 1977), pornography (Pierce, 2001), drug use (Kessler, 2006), masturbation (Kellogg, 1890), genetically modified foods (Druker, 2015), and political correctness (Lukianoff & Haidt, 2015).

These perceptions of harm are not just post-hoc rhetoric, as people see suffering in the faces of children milliseconds after being primed with ostensibly harmless moral violations (e.g. masturbation, Gray et al., 2014). Certainly people can reason about harm after the fact, and do use the rhetoric of harm to convince others (Kahan, 2007; Sood & Darley, 2007), but this deliberative reasoning likely stems from intuitive and automatic perceptions of harm (Kruglanski & Gigerenzer, 2011). At the very least, it is a logical fallacy to conclude that, because perceptions of harm can be used in reasoning, all harm is necessarily reasoned.

Consider breathing as an analogy. No one doubts that people can consciously and deliberatively control their breathing, but this doesn't mean that all breathing is conscious and effortful. In fact, breathing is seldom deliberative. Likewise, the importance of harm in moral justification doesn't imply that perceived harm is always or even frequently a product of justification. One could argue that ubiquitous presence of harm in explicit moral dialogue actually supports its intuitive punch. Consistent with this line of thought, dyadic morality suggests that harm is important and intuitive in moral cognition, and also that most acts are

judged as wrong to the extent that they appear harmful—including disgusting but ostensibly harmless acts.

Disgust and harm

Arguments for the immorality of disgusting acts are often accompanied by the language of harm. Leon Kass claimed that the practice of cloning was not only repugnant, but also intrinsically harmful (Kass, 1997). Likewise, Paul Cameron claimed that the practice of homosexuality was not only disgusting, but also led directly to the sexual molestation of children and the spread of diseases (Nussbaum, 2010). Other theories of morality consider this perceived harm epiphenomenal and post-hoc (Haidt & Hersh, 2001; Haidt, 2012), but recent research suggests otherwise. Consistent with dyadic morality, experiments show a robust causal effect of harm on moral condemnation for both disgusting and non-disgusting deeds, whether harm is induced via explicit statements (Schein & Gray, 2015) or manipulations of threat (Nail, McGregor, Drinkwater, Steele, & Thompson, 2009). For example, simply telling participants that an act was harmful (i.e., "caused others to suffer either emotionally or physically") leads to robust moral condemnation (Study 1; Schein & Gray, 2015). In contrast to the power of harm, the inducement of incidental disgust impacts moral condemnation either weakly or not at all (Landy & Goodwin, 2015).

We acknowledge that most moral psychologists—even those who endorse basic emotions and oppose dyadic morality (Graham et al., 2013)—would agree that exogenously manipulating harm can impact moral judgments. However, dyadic morality suggests that disgusting acts—even those that are "objectively" harmless—are seen as immoral to the extent that they *intrinsically* engender perceptions of harm (i.e., seem inherently harmful). This may seem to be a contradiction in terms (inherent harm in harmless deeds?), but recall that harm is a matter of

perception. Just because researchers deem a disgusting act to be harmless doesn't mean that participants do too (Shweder, 2012). Consistent with this line of thought, past studies finds that people do indeed see harm in ostensibly harmless disgusting acts, and that its perception is more important than ratings of "impurity"—for both liberals and conservatives (Schein & Gray, 2015).

Despite revealing the perceived presence of harm in disgusting acts, these past studies stop short of revealing harm's causal importance within the moral condemnation of disgusting acts. The current studies attempt to address this void by testing mediation patterns among disgust, harm and moral condemnation. If harm is the "active ingredient" in the moral condemnation of ostensibly harmless disgusting acts, then it should mediate the effect of disgust on moral condemnation.

The current research

In three studies, we used mediation analyses to examine links between ratings of disgust, perceived harm, and immorality for ostensibly harmless acts—gay marriage, sacrilege and various sex acts (e.g., anal sex). Notably, these acts have been used by other researchers to argue *for* the power of disgust and *against* the importance of harm in moral judgment (Haidt, 2001; Haidt, Koller, & Dias, 1993; Inbar et al., 2012, 2009). We acknowledge that regression analyses do not yield strict causal evidence, but neither do previous studies that contrast the role of disgust and harm (Haidt & Hersh, 2001; Horberg et al., 2009; Landy & Goodwin, 2015)—even though some of them have made strong causal claims (Haidt & Hersh, 2001).[2] Nevertheless, certain statistical patterns are more or less consistent with the competing hypotheses of *direct disgust* versus *dyadic morality*, as each makes a different prediction about the most proximate predictor of moral judgment.

The historically popular direct disgust hypothesis suggests that moral judgment is most proximately predicted by "disgust and discomfort, which are later cloaked by harm-based rationalizations" (Haidt & Hersh, 2001, p. 212). It hypothesizes that disgust should predict moral condemnation better than perceptions of harm, especially for the ostensibly harmless acts we examine here (Haidt, 2001). More specifically, direct disgust hypothesizes a mediation pattern in which the direct link between disgust and morality should be largely unaffected by adding or removing harm to the model, because harm is ostensibly epiphenomenal and—statistically speaking—should "follow" the moral judgment. See Figure 1.

Dyadic morality instead hypothesizes that moral judgment is most proximately predicted by endogenously perceived harm—even for ostensibly harmless "purity" acts. More specifically, dyadic morality hypothesizes a mediation pattern in which the effect of disgust on morality is mediated by perceived harm, such that any direct link between disgust and moral judgment is decreased or eliminated once accounting for perceptions of harm. In other words, dyadic morality suggests that perceived harm is—statistically speaking—the most important "active ingredient" in the link between disgust and immorality. See Figure 1.

Study 1: The Threat of Gay Marriage

As an initial investigation into links between disgust, harm, and immorality, this study attempts to replicate a well-cited study linking individual differences in disgust with moral condemnation of an ostensibly harmless act (i.e., moral opposition to gay marriage; Inbar et al., 2009) To assess individual differences in perceived harm, we use the Belief in a Dangerous World scale (Altemeyer, 1988). This measure may not seem to be the most obvious choice for testing perceived harm, but danger is a central component of harm, and this well-validated scale has been frequently used in moral and political psychology (van Leeuwen & Park, 2009;

Thórisdóttir & Jost, 2011; Wright & Baril, 2013). Much of this research has shown a

correspondence between BDW and right wing authoritarianism, which in turn predicts increased

moralization (Kugler, Jost, & Noorbaloochi, 2014)—explaining the apparently broader moral

sphere of conservatives (Haidt, 2012). If conservatives see more harm in the world, they should

therefore see more potential for immorality. Given this link, we controlled for political

orientation in all models.

In line with past work, we expect disgust to predict moral condemnation of gay marriage

(Inbar et al., 2009). Importantly, we also predict that perceived harm will mediate the link

between disgust and moral condemnation, consistent with dyadic morality.

Method

Participants

A power analysis suggested that 78 participants would be needed to detect medium effect

sizes ($d = .39$) for both *X* on *M* and *M* on *Y* with a power of .8 using the bootstrap method of

mediation (Fritz & MacKinnon, 2007). As this is a replication, and the original study used a

much larger sample size, we doubled the number of participants suggested by the power analysis.

United States participants with HIT approval rate over 90% completed the survey online through

Amazon Mechanical Turk (mTurk) and were paid $0.20. Of the 187 participants who completed

the study online, 18 failed an information attention check, leaving 169 participants (50% male,

59% liberal, $M_{age} = 35$).

Procedure

All participants completed the survey in one of two orders—with perceived harm

measured either before or after questions regarding attitudes towards gay individuals and

marriage. Order did not influence any of the results, so was dropped from analysis. Participants

completed three measures regarding attitude toward gay individuals and marriage: an explicit measure of attitudes toward lesbian and gay individuals measured using feeling thermometers (Inbar et al., 2012), an explicit measure regarding judgments of same-sex marriage, and an intuitive measure of attitudes (Inbar et al., 2009). Since both explicit measures showed similar patterns (perceived harm fully mediated the link between disgust sensitivity and negative attitudes), and since our focus is on moral disapproval and not explicit attitudes of gay individuals more broadly, results for the feeling thermometer measure are described in *supplementary online materials*. As in Study 1 in Inbar et. al. (2009), participants then completed a disgust sensitivity measure and demographics information.

Measures

Perceived Harm. Altemeyer's (1988) Belief in a Dangerous World Scale was used as a measure of individual differences in perceived harm (sample items: "any day now, chaos and anarchy could erupt around us. All the signs are pointing to it," "our country is not falling apart or rotting from within"). Other scholars (Thórisdóttir & Jost, 2011) have noted that one of the questions in the BDW has strong apocalyptic, and thus religious overtones ("the end is near"). To avoid potential confounds with religiosity, we ran all analysis both with and without this one question. Since these analyses resulted in similar findings, we report only the analysis with the full BDW ($\alpha = .94$).

Disgust Sensitivity. Participants completed the Disgust Sensitivity Scale-Short form (Haidt, McCauley, & Rozin, 1994). Following Inbar et al. (2009), questions regarding sex were excluded from analysis because they are *a priori* conceptually related to sexual morality. The reliability of the six questions used in this study is relatively low, ($\alpha = .67$), but we use this scale to be consistent with past work.

Moral Condemnation of Same-Sex Marriage. Participants answered four questions regarding same-sex marriage (α = .95). Participants rated whether same-sex marriage should be legal and whether states should only recognize marriages between men and women from 1 (*definitely not*) to 5 (*definitely yes*). These items were reversed score, so higher values indicate more negative attitudes. Participants then rated the immorality and wrongness of same-sex marriage from 1 (*not immoral/wrong at all*) to 5 (*extremely immoral/wrong*).

Implicit Attitudes. Inbar and colleague's (2009) adaption of the Knobe effect (Knobe, 2006) was used as a measure of implicit attitudes toward gay intimacy. Participants read that a musical director depicted a couple kissing in his video, and the video had the effect of encouraging public kissing. Half of the participants read that the video depicted gay men, whereas the other participants simply read that it was a couple. The key dependent variable was responses to the question, "did the director intentionally encourage homosexual men to French-kiss in public?" rated on a scale from 1 (*Not at All*) to 7 (*Definitely*). Since negative acts are rated more intentional (Knobe 2006), an increase rating in intentionality for the gay kissing situation would indicate greater disapproval of gay intimacy. As a validation of this measure, the intentionality judgment for gay men kissing significantly correlated with both the explicit ratings on the feeling thermometer, $r(85)$ = -.37, $p < .001$, 95% CI: [.17, .54] and the explicit condemnation of same-sex marriage, $r(85)$ = .51, $p < .001$, 95% CI: [.33, .65]. In contrast, the ratings of the intentionality of the straight people kissing did not significantly correlate with negative attitudes about LGBT individuals, $r(80)$ = - .002, p = .96, 95% CI: [-.22, .22], or gay marriage, $r(80)$ = .05, p = .66, [-.17, .27].

<center>Results</center>

Moral Condemnation of Same-Sex Marriage

Zero order correlations. Consistent with past research (Inbar et al., 2009), disgust sensitivity was significantly correlated with moral condemnation of same-sex marriage, $r(167) =$ .33, $p < .001$, 95% CI: [.19, .46]. Consistent with dyadic morality, perceived harm was significantly correlated with moral condemnation, $r(167) = .60, p < .001$, 95% CI: [.49, .69]— even for this ostensibly harmless act.

Regression analysis. Perceived harm and disgust sensitivity were both standardized and entered into a linear regression model predicting moral condemnation of same-sex marriage. Since this study is looking specifically at individual differences, we controlled for age and politics in the model. Perceived harm significantly predicted moral condemnation of same-sex marriage, $b = .54, SE = .09, t(164) = 5.84, p < .001$, 95% CI: [.36, .72]. Political affiliation was also a significant predictor—more conservative, more immoral—$b = .55, SE = .08, t(164) = 6.62,$ $p < .001$, 95% CI: [.38, .71], though age was not, $b = .009, SE = .006, t(164) = 1.43, p = .15,$ 95% CI: [-.003, .02]. Disgust sensitivity was not a significant predictor of moral condemnation when perceived harm was included in the model, $b = .09, SE = .09, t(164) = 1.01, p = .32,$ 95% CI: [-.08, .25].

Mediation analyses. Perceived harm fully mediated the link between disgust sensitivity and moral condemnation of same-sex marriage, even when controlling for politics and age (Figure 2). There was a significant total effect of disgust sensitivity on moral condemnation, $b = .32, SE = .08, t(164) = 3.82, p = .0002,$ 95% CI: [.15, .48]. Disgust sensitivity was associated with increased perceived harm, $b = .43, SE = .06, t(165) = 6.71, p < .001,$ 95% CI: [.30, .55], and perceived harm was associated with increased moral condemnation, $b= .54, SE =.09, t(164) =$ 5.84, $p < .001$, 95% CI: [.36, .75]. PROCESS (Hayes, 2012) was used to calculate the indirect effect of disgust sensitivity on moral condemnation through perceived harm. The indirect effect

was estimated to be .23, SE = .07, 95% CI: [.12, .38], with 10,000 bootstrap samples, leaving a non-significant direct effect, b = .09, SE = .09, $t(164)$ = 1.01, p = .32, 95% CI: [-.08, .25]. The proportion of mediated effect (PME; indirect/total), was .72 suggesting that this pathway accounts for a high proportion of the effect of disgust on increased moral condemnation. The reverse mediation path was not significant, indirect effect: .04, SE =.05, 95% CI: [-.05, .14], with 10,000 bootstrap samples.

Implicit Attitudes

Zero order Correlations. Implicit disapproval of gay men kissing significantly correlated with disgust sensitivity, $r(85)$ = .29, p = .006, [.08, .47] and perceived harm, $r(85)$ = .44, p < .001, 95% CI: [.25, .60].

Regression analysis. Perceived harm and disgust sensitivity were both standardized and entered into a regression model controlling for politics and age. Perceived harm significantly predicted increased perception of intentionality, b = .58, SE = .22, $t(82)$ = 2.66, p = .009, 95% CI: [.15, 1.01]. Politics was also a significant predictor, b = .40, SE = .19, $t(82)$ = 2.17, p = .03, 95% CI: [.03, .77], although age was not, b = .02, SE = .02, $t(82)$ = 1.25, p = .21, 95% CI: [-.01, .06]. Disgust was not a significant predictor of implicit attitudes with perceived harm in the model, b = .17, SE = .21, $t(82)$ = .82, p = .42, 95% CI: [-.24, .58].

Mediation Analyses. Perceived harm fully mediated the link between disgust sensitivity and intuitive disapproval of gay kissing, even when controlling for politics, and age (Figure 3). The total effect of disgust sensitivity was b = .45, SE = .19, $t(83)$ = 2.42, p = .02, 95% CI: [.08, .82]. Disgust sensitivity was associated with increased harm sensitivity, b = .48, SE = .09, $t(83)$ = 5.32, p < .001, 95% CI: [.30, .66], and harm sensitivity was associated with higher intentionality ratings b= .58, SE =.22 $t(82)$ = 2.66, p = .009, 95% CI: [.15, 1.00]. The indirect effect was

estimated to be .28, $SE = .12$, 95% CI: [.05, .51], with 10,000 bootstrap samples, leaving a non-significant direct effect, $b = .17$, $SE = .21$, $t(83) = .99$, $p = .41$, 95% CI: [-.24, .58]. The proportion of mediated effect was .62 suggesting that this pathway accounts for a high proportion of the effect of disgust on increased moral condemnation. The reverse mediation path was not significant, indirect effect: .09, $SE = .12$, 95% CI: [-.11, .37], with 10,000 bootstrap samples.

Discussion

These results suggest that perceived harm mediates the previously revealed link between disgust sensitivity and the moral condemnation of same sex-marriage. While not causal, these findings are more consistent with dyadic morality than with the direct disgust hypothesis. Past research has also already revealed that exogenous manipulations of harm causally increase moral disapproval of gay rights (Nail et al., 2009). In the current debates about morality, we suggest that these mediation analyses are even more striking because they involve only the harm perceived endogenously within disgusting but ostensibly harmless acts.

Study 2: The Harm of Sacrilegious Ideas

The importance of disgust has been highlighted not only for sexual violations but also for sanctity violations, to which considerations of harm are argued to be less relevant (Rottman & Young, 2014; Rottman, Kelemen, & Young, 2014; but see Gray, 2014). Past research has shown that exposure to sacrilegious statements (e.g. there is no God) elicits disgust (Ritter & Preston, 2011; but see Royzman, Atanasov, Landy, Parks, & Gepty, 2014). Despite the "pathogen-free" and "victimless" nature of these statements, dyadic morality hypothesizes that people can perceive them as harmful, and these perceptions of harm should most proximally predict ratings of the statement's immorality.

Method

Participants

As in Study 1, a power analysis suggested that 78 participants would be needed (Fritz & MacKinnon, 2007) and in the spirit of replicable science, we doubled this suggested sample size. One hundred forty five participants (57% male, 41% female, 2% not reporting; $M_{age} = 35$, all from the United States) completed the study through Amazon's Mechanical Turk (mTurk).

Procedure

Participants were randomly assigned to read one of 10 religious ideological statements designed to preclude any obvious harm. After contemplating the statement for a moment, participants were asked to indicate their emotional responses. Participants were then asked to provide judgments of how harmful and how immoral they perceived each statement to be. Finally, demographic information including religious affiliation was provided at the end of the survey.

Materials and Measures

Ideological statements. Ten religiously-oriented but harmless statements were constructed. Half were anti-religious (i.e., sacrilegious; "God does not exist," "God is not necessary to explain the origin of the universe,") and half were pro-religious ("Jesus is the way, the truth, and the life," "God is the creator of the universe.") See *Appendix A*. Most statements were relevant to all the major monotheistic religions (Christianity, Judaism, Islam), though they did have a slight tilt toward Christianity, and so our models took into account whether a person self-identified as any denomination of Christianity.

Ratings of Disgust and other Emotions. Participants indicated how much they experienced each of the following emotions in response to reading the statement using 5-point

scales ranging from 1 (*not felt at all*) to 5 (*very strongly felt*): disgust, surprise, happiness/joy, anger, fear/anxiety, and sadness.

Perceived harm. Perceived harm was measured by the composite of three items asking how threatening, dangerous, and harmful the statement was perceived to be ($\alpha = .87$), using respective 7-point scales ranging from 0 (*not at all*) to 6 (*extremely*).

Moral condemnation. Moral condemnation was measured as the composite of three items asking how morally wrong, blameworthy, and immoral the statement was perceived to be ($\alpha = .93$), using respective 7-point scales ranging from 0 (*not at all*) to 6 (*extremely*).

Results

Zero order correlations. Consistent with accounts linking disgust and moral condemnation within pathogen-free religious violations (Ritter & Preston, 2011; Haidt, 2012), feelings of disgust predicted moral condemnation of sacrilegious statements, $r(143) = .45$, $p < .001$ [.31, .57]. Consistent with predictions from dyadic morality, perceived harm also predicted moral condemnation of these statements, $r(143) = .77$, $p < .001$, [.69, .83].

Regression analysis. Perceived harm and disgust were both standardized and then entered into a regression model that predicted moral judgment controlling for whether a person identified as Christian. Harm was the only significant predictor of immorality, $b = 1.28$, $SE = .11$, $t(142) = 11.61$, $p < .001$, [1.06, 1.50]. With harm in the model, disgust was no longer a significant predictor of immorality, $b = .17$, $SE = .11$, $t(142) = 1.54$, $p = .13$, [-.05, .39].

Mediation analyses. Perceived harm fully mediated the link between disgust and moral condemnation, even when controlling for religious affiliation (Figure 4). We calculated the indirect effect of disgust on moral judgment through harm controlling for religion using PROCESS (Hayes, 2012). The data were structured hierarchically (i.e., participants were nested

under one of 10 different stimuli). Since there were not enough clusters to run a multi-level model, we used the fixed-effects approach to clustering (Hayes 2013, p. 11) and partialled out statement-level effects by including the statement as a cluster variable in the model. As seen in Figure 4, the total effect of disgust on moral judgment was statistically significant, $b = .72$, $SE = .14$, $t(133) = 5.14$, $p < .001$, 95% CI: [.44, .99]. Disgust was associated with increased harm $b = 0.42$, $SE = .07$, $t(133) = 5.66$, $p < .001$, 95% CI: [.27, .57], and harm was associated with more severe moral judgments $b = 1.24$, $SE = .12$, $t(132) = 10.17$, $p < .001$, 95% CI: [1.01, 1.48]. The indirect effect was estimated to be .52, $SE = .12$, 95% CI: [.30, .77], with 10,000 bootstrap samples, leaving a non-significant direct effect, $b = .19$, $SE = .12$, $t(132) = 1.66$, $p = .10$, 95% CI: [-.04, .42]. The proportion of mediated effect is .72. The reverse mediation was not significant, indirect effect: .09, $SE = .06$, 95% CI: [-.01, .25], with 10,000 bootstrap samples.

Anger: Exploratory data analysis. In this paper, we are focused on the potential mediating role of harm but also of interest is whether there is a unique link between disgust and purity (Cameron et al., 2015). Since this study collected ratings of anger, another high arousal negative emotion, as exploratory data analysis, we tested whether anger also mediated the previously found link between disgust and moral condemnation. Disgust was associated with increased anger, $b = .52$, $SE = .08$, $t(133) = 6.96$, $p < .001$, 95% CI: [.38, .67], and anger was associated with more moral condemnation, $b = .86$, $SE = .14$, $t(132) = 6.08$, $p < .001$, 95% CI: [.58, 1.15]. Anger fully mediated the effect of disgust on immorality, indirect effect: .45, $SE = .11$, 95% CI: [.26, .71], with 10,000 bootstrap samples, leaving only a marginally significant direct effect of disgust, $b = .26$, $t(133) = 1.82$, $p = .07$. Disgust did not mediate the link between anger and moral condemnation, indirect effect: .13, $SE = .09$, 95% CI: [-.01, .36], with 10,000 bootstrap samples.

Discussion

Consistent with some past research, we found links between disgust and moral condemnation of pathogen-free "sanctity" violations (Ritter & Preston, 2011). Importantly, this connection was mediated by perception of the harmfulness of those statements, consistent with dyadic morality. This disgust-immorality link was also mediated by feelings of anger, consistent with other recent work on the expanded scope of anger across moral diversity (Royzman et al., 2014). For readers that lack religious conviction, these perceptions of harm may seem absurd, but many people link bad thoughts to harmful actions in the world (Shafran, Thordarson, & Rachman, 1996), and so simply entertaining sacrilegious thoughts may seem to have harmful effects. Consider the feeling you get when considering the statement "child pornography is great fun"—even if you know that thoughts are harmless, it may be hard to shake the feeling that seriously considering this statement is dangerous.

Study 3: The Harm of Diverse Disgusting Acts

The previous two studies revealed that perceived harm mediates the link between disgust and the moral condemnation of ostensibly harmless violations. However, these studies used a relatively narrow assessment of disgusting acts. In the current study, we use a more diverse sample of acts drawn from three descriptively different "domains" of disgust (Tybur, Lieberman, et al., 2009): sexual disgust, canonical pathogen disgust, and canonical moral disgust (i.e., assessed via obviously harmful/dyadic acts). We again predict that these disgusting acts will be seen as immoral to the extent that they are seen as harmful.

This study also examines the role that atypicality or weirdness plays in the moral condemnation of disgusting acts. Moral violations are counter-normative (Nichols, 2002) and studies reveal that general counter-normativity (i.e., atypicality/weirdness) is an important non-

specific driver of moral judgments (Gray & Keeney, 2015a, 2015b). As many ostensibly

harmless "purity" acts are quite strange (Gray & Keeney, 2015b), we examine here whether

perceptions of weirdness might also help mediate links between disgust and immorality.

<center>Method</center>

Participants

A power analysis for the indirect effect of a multi-level multiple mediation model

computed with a Sobel test requires *a priori* estimations of complicated covariance matrixes and

advanced simulation techniques, which are themselves imperfect given the number of parameters

they need to take into consideration (e.g. interclass correlation coefficients, each path on the

mediation). As a rough estimation of power, we conducted a power analysis for typical

mediation models using the Sobel test. To detect medium effects with power set at .8, 90

participants are needed (Fritz & MacKinnon, 2007). Nevertheless, due to the difficulty of

determining the correct power, we opted for substantially more participants. Of the 212

participants who completed the study online, 43 failed to complete all questions or failed an

information attention check, leaving 169 participants (56% female, 44% male, 1 participant not

identifying according to the gender binary, 59% liberal, $M_{age} = 35$). As this is a multi-level model

exploring Level 1 variance, the Ns used in analyses are effectively much higher.

Procedure

Participants rated the perceived harmfulness, disgustingness, weirdness, unpleasantness

and immorality of 24 different disgusting actions. The order of the questions, and the order of the

actions listed within each question block were both randomized. At the end of the survey,

participants provided demographic information.

Materials

Stimuli. The 24 acts, taken directly from Tybur et al. (2009) included eight sexually disgusting acts (e.g. performing oral sex, having sex in exchange for money), eight pathogen-related disgusting actions (e.g. seeing a cockroach run across the floor, finding a hair in your food), and eight canonically harmful moral violations (stealing from a neighbor, deceiving a friend). Past research has suggested that all these acts evoke disgust (Tybur, Lieberman, et al., 2009).

Ratings. As in past research (Gray & Keeney, 2015b), three questions assessed perceived harm (harmful, threatening, dangerous; α = .87), disgust (gross, disgusting, unnatural; α = .75), and weirdness (weird, bizarre, atypical; α =.74), and two questions measured immorality (immoral, wrong; α =.89), and one question measured perceived unpleasantness. All questions were answered on a 6-point scales ranging from 1 (*not at all*) to 6 (*extremely*).

Results

For completeness, we present two sets of analyses. The first examines the role of perceived disgust and harm for all disgusting acts within the Tybur scale (Tybur, Lieberman, et al., 2009). Although this is the most comprehensive analysis, it also includes canonically immoral acts (i.e., harmful acts). Therefore, the second analysis removes these canonically immoral acts from the analysis.

All acts

Zero order correlations. All rating categories were significantly correlated with moral condemnation: perceived harm, $r(4054) = .74$, , $p < .001$, 95% CI: [.73, .75] being the strongest predictor followed by perceived weirdness, $r(4054) = .45$, $p < .001$, 95% CI: [.43, .47], unpleasantness $r(4056) = .36$, $p < .001$, 95% CI: [.33, .39], and disgust, $r(4054) = .36$, $p < .001$, 95% CI: [.33, .39].

Regression analysis. Perceived harm, disgust, weirdness and unpleasantness were all entered as Level 1 predictors into a multi-level model predicting moral condemnation. All factors were person mean centered to assess within subject variance and control for differences between people. Notably, perceived harm significantly predicted immorality, even when controlling for all other factors $B = .94$, $SE= .02$, $t(3883) = 59.09$, $p < .001$, 95% CI: [.91, .97]. Perceived weirdness also predicted immorality, $B = .22$, $SE= .02$, $t(3883) = 9.52$, $p < .001$, 95% CI: [.17, .26]. Although zero-order correlations found that perceived disgust positively predicted immorality, once controlling for other factors, perceived disgust had a *negative* relationship with immorality, $B = -.26$, $SE= .02$, $t(3883) = -10.57$, $p < .001$, 95% CI: [-.31, -.22]. Across all acts—controlling for perceived harm—the more an act is perceived as disgusting, the *less* immoral the act is judged—likely because of the inclusion of prototypically immoral (i.e., harmful) acts. Perceived unpleasantness was not a significant predictor of immorality, $B = .03$, $SE= .02$, $t(3883) = 1.57$, $p = .12$, 95% CI: [-.01, .06], so was dropped from all further analyses.

Mediation analyses. Adapting the guidelines by Zhang, Zyphur, & Preacher (2009), a multilevel multiple mediation model using the Sobel method was run to test whether harm mediated the link between disgust and immorality, and whether weirdness was also a mediator. As seen in Figure 5, there was a significant total effect of disgust on immorality, $B = 0.37$, $t(3886) = 15.30$, $SE = .02$, $p < .001$, 95% CI: [.32, .41]. Disgust was associated with increased harm, $B = 0.50$, $SE = .02$, $t(3886) = 28.84$, $p < .001$, 95% CI: [.47, .54] and harm was associated with more severe moral judgments, $B = .95$, $SE = .02$, $t(3884) = 59.80$, $p < .001$, 95% CI: [.91, .98]. Disgust was also associated with perceived weirdness, $B = .61$, $SE = .01$, $t(3886) = 50.18$, $p < .001$, 95% CI: [.58, .63], and perceived atypicality/weirdness had a significant association with moral condemnation, $B = .22$, $SE = .02$, $t(3884) = 9.67$, $p < .001$, 95% CI: [.18. .27]. The

indirect effect of perceived disgust through perceived harm controlling for the indirect effect of weirdness is .48, *SE* = .0003, *Z* = 1250.00, *p* < .001, and the indirect effect of weird controlling for the indirect effect of harm is .13, *SE* = .00004, *Z* = 3050.00, *p* < .001. After accounting for perceived harm and weirdness, there was a negative direct effect of disgust, *B* = -.24, SE = .02, *t*(3884) = -11.50, *p* < .001, 95% CI: [-.28, -.20]. The proportion of mediated effect through harm was 1, and the proportion of weirdness is .36, suggesting that the harm pathway accounts for a large part of the connection between disgust and increased moral condemnation. The reverse mediation effect (perceived disgust mediating the link between perceived harm and immorality, controlling for the indirect effect of weirdness) was not significant.

Across all acts, disgust predicted increased moral condemnation in as much as the act was also perceived as harmful. Once accounting for the shared variance with harm, disgust negatively predicted immorality.

Removing the harmful, canonically immoral acts

Dyadic morality predicts that the immorality of even ostensibly harmless disgusting acts should be predicted by perceived harm. We therefore reran the analysis excluding the canonically harmful immoral acts.

Zero order correlations. All rating categories significantly correlated with immorality of ostensibly harmless disgusting acts: Harm, *r*(2702) = .69, *p* < .001, 95% CI: [.67, .71], and disgust *r*(2702) = .59, *p* < .001, 95% CI: [.56, .61], had the strongest correlation followed by weirdness, *r*(2702) = .59, *p* < .001, 95% CI: [.56, .61], and then unpleasantness, *r*(2702) = .44, *p* < .001, 95% CI: [.41, .47].

Regression analysis. Perceived harm, disgust and weirdness were all entered as Level 1 predictors into a multi-level model predicting immorality. Of note, the 95% confidence intervals

suggest that perceived harm, $B = .48$, SE $= .02$, $t(2532) = 26.23$, $p < .001$, 95% CI: [.44, .52], is

a (much) stronger predictor of immorality than perceived disgust, $B = .05$, SE $= .02$, $t(2532) = $

2.37, $p = .02$, 95% CI: [.009, .09], even when controlling for weirdness, $B = .28$, SE $= .02$,

$t(2532) = 13.73$, $p < .001$, 95% CI: [.24. .32].

Mediation analyses. Once again, a multilevel multiple-mediation model using the Sobel

method was run to test whether harm or weirdness mediated the link between disgust and

immorality (Figure 6). There was a significant total effect of perceived disgust, $B = 0.51$, SE $=$

.02, $t(2534) = 28.27$, $p < .001$, 95% CI: [.47, .54]. Perceived disgust was associated with

increases in perceived harm, $B = 0.57$, SE $= .02$, $t(2534) = 34.52$, $p < .001$, 95% CI: [.54, .61],

and perceived harm, controlling for perceived weirdness, was associated with more severe moral

judgments $B = .48$, SE $= .02$, $t(2532) = 26.23$, $p < .001$, 95% CI: [.44, .52]. Disgust was also

associated with perceived weirdness, $B = .64$, SE $= .01$, $t(2534) = 42.89$, $p < .001$, 95% CI: [.61,

.67], and perceived weirdness had a significant association with moral condemnation, $B = .28$,

SE $= .02$, $t(2532) = 13.73$, $p < .001$, 95% CI: [.24. .32]. The indirect effect of perceived disgust

through perceived harm controlling for the indirect effect of weirdness is .27, $SE = .0002$, $Z =$

1425.00, $p < .001$, and the indirect effect of weird controlling for the indirect effect of harm is

.18, $SE = .00004$, $Z = 3200.00$, $p < .001$. After accounting for perceived harm and weirdness,

there was a small direct effect of disgust, $B = .05$, SE $= .02$, $t(2532) = 2.37$, $p = .02$, 95% CI:

[.009, .09]. The proportion of mediated effect through harm was .54, and the proportion of

weirdness is .35, suggesting that the perceived harm and weirdness pathways accounts for a large

part of the connection between disgust and increased moral condemnation.

Given the large amount of power ($df = 2532$), the reverse mediation effect—perceived

disgust mediating the link between perceived harm and immorality controlling for perceived

weirdness—was significant, $p < .001$, though its effect size was extremely small, $B = .02$, SE =.00002.

Discussion

Within a diverse sample of disgusting acts, it appears that moral condemnation is most proximally predicted by perceived harm. Importantly, this mediation effect holds whether we examined all disgusting acts or only ostensibly harmless "purity" acts (Graham et al., 2013).

General Discussion

Across three studies, perceptions of harm statistically mediated the link between disgust and moral condemnation. Individual differences in perceived harm mediated the link between disgust sensitivity and the moral condemnation of gay intimacy and gay marriage (Study 1). Perceptions of harm mediated the link between felt disgust and the moral condemnation of sacrilegious statements (Study 2). Finally, perceptions of harm mediated the link between ratings of disgustingness and the moral condemnation of a diverse set of disgusting acts (Study 3). Disgust appears to be linked to moral condemnation to the extent that it involves perceived harm, a finding that is more consistent with dyadic morality (Gray & Schein, 2012) than with the historically popular "direct disgust" hypothesis (Haidt & Hersh, 2001; Schnall et al., 2008).

These results both affirm and challenge past work on disgust and morality. Within the subset of acts we investigated, feelings of disgust were robustly related to moral condemnation (replicating Inbar et al., 2009). However, this relation was not special to disgust. Instead, it was statistically accounted for by perceived harm, which transcends any specific emotions to provide a broad cognitive template for immorality (Royzman, Goodwin, & Leeman, 2011; Schein & Gray, 2015). The role of harm may therefore be a "hidden mediator" to help explain conflicting reports about the ability for incidental disgust to impact moral judgment (Landy & Goodwin,

2015). Perhaps incidental disgust amplifies moral judgment to the extent that any specific operationalization of disgust simultaneously primes danger, threat, or vulnerable victims.

Caveats

Any set of studies comes with caveats and the current studies are no exception. First, there is broad diversity among both immoral and disgusting acts (Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011), and the current studies examined a relatively small subset of these acts. However, we selected these violations—gay marriage, religious blasphemy and sex acts—specifically because they have been used in past research to argue for the irrelevance of harm in the moral condemnation of disgusting acts (Haidt & Hersh, 2001). Thus, they represented the most conservative test of our hypothesis. Study 3 also investigated a broad scope of disgusting acts as determined by an impartial and well-validated scale of disgust (Tybur, Lieberman, et al., 2009).

Second, we acknowledge that all participants were Americans, and research has revealed that their moral judgments may be relatively "WEIRD" (Barrett, 2013; Lindquist, 2013). However, past work on disgust and "purity" (Inbar et al., 2012, 2009) draws from the same population (Graham et al., 2013) and research by both us and others reveals that non-Americans and American non-liberals see harm within ostensibly harmless purity violations (Gray et al., 2014; Schein & Gray, 2015; Shweder et al., 1997). Interestingly, one recent study does seems to argue against the sweeping role of harm in morality across cultures (Buchtel et al., 2015), as it finds that Chinese people view acts of incivility as more "immoral" than acts of harm. However, it is unclear whether the authors' Chinese translation of "morality" captures the same concept in the West, as the authors report that Chinese participants see murdering an elderly woman as much *less* immoral than failing to give her your bus seat (Study 3; Buchtel et al., 2015). As

morality is tied to behavior—and presumably there are fewer Chinese murderers than selfish

bus-sitters—more cross-cultural research is surely needed.

Third, we acknowledge that mediation analyses do not reveal strict causal evidence.

Nevertheless, these mediation patterns are certainly much more consistent with dyadic morality

than with the direct disgust hypothesis. Also, while it is true that correlation does not imply

causation, the *absence* of correlation certainly implies the *absence* of causation. Here, the

reliably absent correlation between disgust and immorality in our mediation models (not to

mention in meta-analyses; Landy & Goodwin, 2015) argues against the direct disgust model. The

causal role of harm is also supported by multiple experiments revealing that manipulating

intention, causation and/or damage (i.e., all elements of dyadic harm) robustly alters moral

judgment (Cushman, 2008; Cushman & Young, 2011; Guglielmo, Monroe, & Malle, 2009;

Sousa, Holbrook, & Piazza, 2009).

Finally, one limitation of these studies is that they rely upon explicit self-reports, which

can be biased by lay-theories and post-hoc justification. However, there seems to be little need

for our anonymous participants to justify themselves in these studies. Our past research also

reveals that participants automatically and implicitly see harm in disgusting and "harmless"

violations (Gray et al., 2014). These self-report judgments also appear to be valid enough to be

used extensively by disgust researchers (Haidt et al., 1994; Tybur, Lieberman, et al., 2009), even

those advocating for the direct disgust hypothesis (e.g. Studies 1 & 3, Horberg et al., 2009).[3]

Moreover, as we mentioned in the introduction, the existence of post-hoc harm-based

justifications does not argue against a causal and automatic role for harm. If anything, such

harm-based reasoning actually supports dyadic morality, given the frequent continuity between

intuitive and deliberative judgments (Kruglanski & Gigerenzer, 2011).

Construction versus modularity

These results support the theory of emotion construction by revealing the importance of a more basic ingredient within disgust—perceived harm. We acknowledge that basic emotion accounts argue for harm's *ultimate*—i.e., evolutionary—role in disgust (Chapman & Anderson, 2012), but the current results reveal a *proximate* psychological role for harm. Importantly, these results reveal that disgust is not a monolithic experience with an invariant effect on moral judgment. Instead it is a collection of experiences with variable perceptions of harm, and it is these perceptions of harm that best predict moral condemnation.

By revealing disgust to be a heterogeneous set of experiences composed of more fundamental ingredients, these results add to evidence (e.g., Cameron et al. 2015) that challenges theories of morality inspired by basic emotions that posit a 1:1 mapping between specific emotions and moral content (Graham et al., 2013; Horberg et al., 2009; Rozin et al., 1999). The most popular of these theories is "moral foundations theory" (MFT) which argues for a whole number of moral mechanisms, or "little switches in the brain of all animals" (Haidt, 2012, p. 123). These mechanisms are argued to be "interestingly" modular—distinct, domain-specific, and specifically tied to one emotion (Haidt & Joseph, 2005)—and to be triggered by a specific kind of moral content (e.g., fairness, loyalty, purity). Receiving the most attention is the "purity," defined as violations of sexual propriety and religious orthodoxy (Feinberg & Willer, 2013; Haidt, 2012; Inbar & Pizarro, 2014; Young & Saxe, 2011). Given that these are the exact kind of violations investigated in these studies, the current data speaks to the idea of a "purity foundation."

What is Purity?

MFT suggests that the "purity foundation" is a mechanism whose operation is distinct from concerns about harm (Haidt, 2012), involves different cognitive operations than harm (Young & Saxe, 2011), and is specially linked to disgust (Graham et al., 2013; Horberg et al., 2009). However, recent work—and the current studies—casts doubt on all three of these modular "foundation" claims. Arguing against distinctness, correlations between ratings of harm and purity are so high as to be indistinguishable ($r > .87$ Gray & Keeney, 2015b), even when using measures developed specifically by MFT (see Gray & Keeney, 2015a for a discussion of the Moral Foundation Questionnaire). Most strikingly, purity appears to fail its own manipulation checks, as canonically harmful acts (e.g., abuse) are seen to be *more* "impure" than so-called impurity violations (e.g., chicken masturbation; Gray & Keeney, 2015b)—likely because some definitions of impurity are synonymous with "immorality" which is best predicted by harm (Schein & Gray, 2015). The current studies also reveal the robust presence of perceived harm in "purity" acts, further arguing against its distinctness from harm.

Arguing against different cognitive operations (i.e., domain-specificity), research reveals that apparent differences between harm and purity (e.g., act/character dissociations; Uhlmann & Zhu, 2013) stem from scenario confounds. The "purity" scenarios of chicken masturbation (Haidt, 2001) and corpse pizza-eating (Clifford, Iyengar, Cabeza, & Sinnott-Armstrong, 2015) are weirder (i.e., less typical) and less severe than the "harm" scenarios of murder and assault (Gray & Keeney, 2015b). Experiments reveal that these confounds in typicality/weirdness and severity lead to cognitive differences, rather than manipulations of purity versus harm *per se* (Gray & Keeney, 2015b). For example, people think that eating pizza off a corpse is very odd, and this oddness—rather than "purity"—is what causes judgments of poor moral character.

Importantly, considerations of severity and typicality are both consistent with a dyadic template for moral judgment.

A recent review of the moral emotion literature also fails to find specific links between disgust and "purity" violations, instead revealing more general affective and conceptual associations between emotion and morality (Cameron et al., 2015), consistent with constructionism. The current research also reveals that disgust *per se* is a poor predictor of moral judgment in purity acts, at least once statistically accounting for perceptions of harm. The predictive power of perceived harm in even ostensibly harmless acts provides a new perspective on the logic behind the "purity mechanism."

The Logic Behind "Purity"

Decades ago, researchers noticed that different cultures condemned acts (e.g., sexual propriety) that seemed harmless to their liberal, American eyes (Haidt et al., 1993). Taking this harmlessness at face-value, they argued against the reigning harm-based theory of morality (Turiel et al., 1987), and for the existence of distinct harm-independent mechanisms (e.g., purity; Haidt & Joseph, 2004). The logic was that if disgusting acts are judged as both harmless and wrong, there must be a harm-independent-related mechanism why disgusting acts seem wrong—hence the purity mechanism. However, the current research suggests that these "harmless" acts are not harmless at all, which therefore obviates the need to posit an additional (and less parsimonious) "purity mechanism" to account for cultural differences. One can account for the immorality of sexual and religious violations simply by their perceived harmfulness.

The key to maintaining both cognitive parsimony and anthropological pluralism is to recognize that harm—like morality—is a matter of perception (Schein & Gray, 2015), rather than an objective, rational fact (Haidt, 2012). Indeed, dyadic morality strongly endorses moral

pluralism (Schein et al., 2015), but through the combinatorial process of constructionism.

Specifically, diverse moral concerns are psychologically instantiated as different varieties of

harm (e.g., physical, social, spiritual) that occurs between different agents and patients (Gray,

Waytz, et al., 2012; Schein et al., 2015). For example, those who see gay marriage as wrong and

"impure" perceive it to cause harm through the damnation of one's future self, the weakening of

societal institutions which encourage order, and the suffering of children (Bryant, 1977).

Clarifying the terms of the debate

The weight of evidence here and elsewhere (Cameron et al., 2015; Gray & Keeney,

2015a, 2015b; Schein & Gray, 2015) suggests that MFT does not describe moral modules in the

strict sense. Accordingly, MFT now emphasizes that it describes only the "first draft" of a moral

mind, an innate evolutionary blueprint that interacts with cultural learning (Haidt, 2012, p. 153;

Haidt, Graham, & Ditto, 2015). This statement, however, is non-specific, as many theories

acknowledge the dual importance of innateness and cultural learning, including dyadic morality

(Govrin, 2014). This claim of "minimal" modularity also directly contradicts recent claims that

MFT is "near the maximalist side of the [modularity] spectrum" (Graham et al., 2013, p. 99).

Despite this contradiction, "minimal modularity" might seem to provide a reasonable

compromise between competing views; unfortunately, we suggest that this intermediate position

is empirically unsound.

To its credit, the strong version of modularity provides many clear testable predictions—

distinctness, domain-specificity, emotion-specificity—which is why past research on dyadic

morality has focused on disconfirming them (Cameron et al., 2015; Gray & Keeney, 2015b;

Schein & Gray, 2015). However, the minimal version of MFT modularity does not provide such

clear predictions. Instead it endorses exactly opposite claims, making it consistent (and also inconsistent) with any set of empirical findings—and therefore unfalsifiable.

Consider the contradiction in defining moral foundations as "little switches in the brain," (Haidt, 2012, p. 123), while also explicitly stating that "foundations are *not* spots in the brain" (Graham et al., 2013, p. 96, italics added). Or suggesting that moral foundations are "different kinds of moral concerns" (Graham et al., 2011, p. 381) but are not "distinct" (Haidt et al., 2015).

MFT has also denies that moral foundations reflect modular "domain-specific" systems (Haidt et al., 2015), while simultaneously arguing for "distinct systems [that] subserve different types of moral judgment" (Parkinson et al. 2011, quoted in Graham et al., 2013, p. 98) that generate "specific emotional and motivational reactions"—all of which are defining characteristics of modular domain-specificity. Even the claim that MFT endorses both "similarities and differences...between different kinds of moral judgment" (Graham, 2015, p. 872) is empirically unresolvable.

Synthesis: Content vs. Process, Naming vs. Explaining

One potential way of synthesizing competing perspectives is through the classic division of content versus process. Constructionist dyadic morality provides a parsimonious *process* for moral cognition, one that is consistent with both intuitionism and pluralism. The dyad functions automatically (consistent with intuitionism; Haidt, 2001) and facilitates diversity (consistent with pluralism; Shweder, 2012) through the combination of different agents and patients in different contexts. However, we acknowledge that dyadic morality does not document which varieties of harm are important across cultures.

As we have suggested before, there is no doubt that sexual chastity, religious purity, economic frugality, military duty and familial caring all represent descriptively different moral

content, whose importance can vary across people, place and time. For example, dyadic morality accepts that moral language can vary across culture in Twitter (Dehghani et al., in press) and in Congress (Graham, Haidt, & Nosek, 2009), but such language does not illuminate underlying moral psychological *processes* (just as cultural differences in music fails to illuminate the underlying processes of auditory cognition). We agree that such content differences can be meaningfully organized by taxonomies (Haidt, 2012; Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011), but these taxonomies—like lists of origami shapes—must not be reified as natural kinds (Barrett, 2009). In morality and elsewhere, our common sense categories seldom represent fundamental mental processes (Nisbett & Wilson, 1977).

It is also important to distinguish "naming" from "explaining" (Gawronski & Bodenhausen, 2015). That liberals and conservatives differ on concerns of sexual and religious morality is uncontroversial given decades of work on right-wing-authoritarianism (Altemeyer, 1988; Kugler et al., 2014). Labeling this a difference of "purity" provides a useful name, but does not *explain* it, as MFT claims (Haidt, 2012). Useful explanation requires grounding phenomena in underlying (i.e., more fundamental) mechanisms (i.e., the "how") rather than simply providing an intuitive label (i.e., the "what"). Consider the explanation behind birds' ability to fly. One may attribute this ability to a "flight capacity" that is both innate and learned, but this label is really a restatement of the phenomena. Instead, the real explanation involves many underlying factors, such as skeletal density, feather structure and early developmental imprinting—"flight capacity" is exactly what we want to understand.

Likewise, explaining cultural differences in "purity" requires more than ascribing it to an eponymous "purity" mechanism. Indeed, if purity is *defined* as sensitivity to sexual and religious morality, then the statement "Conservatives are sensitive to sexual/religious morality because of

their purity foundation" can be reformulated into the tautology, "Conservatives are sensitive to sexual/religious morality because of their sensitivity to sexual/religious morality." Uncovering explanations for cultural differences in morality is an extremely important task, but one that we suggest requires examining deeper individual differences (Kugler et al., 2014), and regional and historical differences in threat or harm (van Leeuwen, Park, Koenig, & Graham, 2012). Researchers may also want to re-examine moral diversity explanations that center on cultural differences in disgust (Haidt et al., 1993) given the current findings about the role of harm.

Conclusion

The language of disgust features prominently in moral condemnation, but not all disgusting acts are immoral. There is no doubt that giving sick children blood contaminated with HIV is both disgusting and immoral (Wheeler, 2015). On the other hand, cleaning up your child's diarrhea is a gross but morally laudable act of parental love. The difference between these deeds lies in harm. Disgusting acts are wrong when they seem to cause harm, even when they are "objectively" harmless. The predictive power of harm not only helps clarify links between disgust and immorality, but also has implications for understanding the nature of emotion and moral cognition. This research is silent on whether feelings of disgust should enjoy normative weight, but to the extent that there is moral "wisdom in repugnance," it is the wisdom of harm.

Acknowledgements

References

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

Fransisco, CA: Jossey-Bass.

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*(1),

28–58.

Barrett, L. F. (2009). The future of psychology: Connecting mind to brain. *Perspectives on

Psychological Science*, *4*(4), 326 –339. http://doi.org/10.1111/j.1745-6924.2009.01134.x

Barrett, L. F. (2013). Psychological Construction: The Darwinian approach to the science of

emotion. *Emotion Review*, *5*(4), 379–389. http://doi.org/10.1177/1754073913489753

Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, *425*(6955),

297–299. http://doi.org/10.1038/nature01963

Bryant, A. (1977). *The Anita Bryant story: The survival of our nation's families and the threat of

militant homosexuality*. Grand Rapids, MI: Revell.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., & Bond, M. H. (2015).

Immorality East and West Are Immoral Behaviors Especially Harmful, or Especially

Uncivilized? *Personality and Social Psychology Bulletin*, *41*(10), 1382–1394.

http://doi.org/10.1177/0146167215595606

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The

psychophysiology of emotion. *Handbook of Emotions*, *2*, 173–191.

Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and

emotions: No evidence for specific links between moral content and discrete emotions.

Personality and Social Psychology Review. http://doi.org/10.1177/1088868314566683

Chapman, H. A., & Anderson, A. K. (2012). Understanding disgust. *Annals of the New York Academy of Sciences*, *1251*(1), 62–76.

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

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

Crawford, J. T., Inbar, Y., & Maloney, V. (2014). Disgust sensitivity selectively predicts attitudes toward groups that threaten (or uphold) traditional sexual morality. *Personality and Individual Differences*, *70*, 218–223. http://doi.org/10.1016/j.paid.2014.07.001

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. http://doi.org/10.1016/j.cognition.2008.03.006

Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science*, *35*(6), 1052–1075. http://doi.org/10.1111/j.1551-6709.2010.01167.x

Dehghani, M., Johnson, K. M., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., … Graham, J. (in press). Purity homophily in social networks. *Journal of Experimental Psychology: General*.

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

Druker, S. (2015). *Altered Genes, Twisted Truth: How the Venture to Genetically Engineer Our Food Has Subverted Science, Corrupted Government, and Systematically Deceived the Public*. Salt Lake City, UT: Clear River Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169–200. http://doi.org/10.1080/02699939208411068

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, *3*(4), 364–370. http://doi.org/10.1177/1754073911410740

Feinberg, M., & Willer, R. (2013). The moral roots of environmental attitudes. *Psychological Science*, *24*(1), 56–62. http://doi.org/10.1177/0956797612449177

Fritz, M. S., & MacKinnon, D. P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychological Science*, *18*(3), 233–239. http://doi.org/10.1111/j.1467-9280.2007.01882.x

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

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion*, *14*(2), 251.

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. (2015). Explaining away differences in moral judgment: Comment on Gray & Keeney (2015). *Social Psychological and Personality Science*.

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

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(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. (2014). Harm concerns predict moral judgments of suicide: Comment on Rottman, Kelemen and Young (2014). *Cognition*, *133*(1), 329–331. http://doi.org/10.1016/j.cognition.2014.06.007

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

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

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

Guglielmo, S., Monroe, A. E., & Malle, B. F. (2009). At the heart of morality lies folk psychology. *Inquiry*, *52*(5), 449–466.

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.

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

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

Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*(1), 191–221. http://doi.org/10.1111/j.1559-1816.2001.tb02489.x

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55–66.

Haidt, J., & Joseph, C. (2005). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind*. New York: Oxford University Press.

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*(4), 613–628.

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences*, *16*(5), 701–713. http://doi.org/10.1016/0191-8869(94)90212-7

Haslam, N. O. (1998). Natural kinds, human kinds, and essentialism. *Social Research*, *65*(2), 291–314.

Hayes, A. F. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling*. Retrieved from http://is.muni.cz/el/1423/podzim2014/PSY704/50497615/hayes_2012_navod_process.pdf

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*(6), 963–976. http://doi.org/10.1037/a0017423

Inbar, Y., & Pizarro, D. (2014). Pollution and purity in moral and political judgment. *Advances in Experimental Moral Psychology*, 111.

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

Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion*, *9*(3), 435–439. http://doi.org/10.1037/a0015960

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

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

Kass, L. (1997, June 2). The wisdom of repugnance. *The New Republic*, *216*(22), 17–26.

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

Kessler, J. G. United States of America v. Philip Morris USA, inc., No. Civil Action No. 99-2496 (United States District Court for the District of Columbia 2006).

Knobe, J. (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies*, *130*(2), 203–231. http://doi.org/10.1007/s11098-004-4510-0

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*(1), 97–109. http://doi.org/10.1037/a0020762

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

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, *10*(4), 518–536. http://doi.org/10.1177/1745691615583128

Lieberman, D., Tooby, J., & Cosmides, L. (2003). Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest. *Proceedings of the Royal Society of London B: Biological Sciences*, *270*(1517), 819–826.

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist model. *Emotion Review*, *5*(4), 356–368.

Lindquist, K. A., & Barrett, L. F. (2008). Constructing Emotion The Experience of Fear as a

Conceptual Act. *Psychological Science*, *19*(9), 898–903.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain

basis of emotion: a meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3),

121–143. http://doi.org/10.1017/S0140525X11000446

Lukianoff, G., & Haidt, J. (2015, September). The Coddling of the American Mind. *The Atlantic*.

Retrieved from http://www.theatlantic.com/magazine/archive/2015/09/the-coddling-of-

the-american-mind/399356/

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

Nail, P. R., McGregor, I., Drinkwater, A. E., Steele, G. M., & Thompson, A. W. (2009). Threat

causes liberals to think like conservatives. *Journal of Experimental Social Psychology*,

45(4), 901–907.

Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment.

Cognition, *84*(2), 221–236.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on

mental processes. *Psychological Review*, *84*(3), 231–259.

Nussbaum, M. C. (2010). *From disgust to humanity: Sexual orientation and constitutional law*.

Oxford University Press. Retrieved from

https://books.google.com/books?hl=en&lr=&id=-

Oq3katPB40C&oi=fnd&pg=PT9&dq=from+disgust+to+humanity&ots=rksQ_C8eBj&si

g=r3IUQcfMmlZAbCx_Rd9rCAv8c4s

Pierce, J. (2001, May 4). Opponents of pornography contend it's not "harmless" or victimless. Retrieved October 16, 2015, from http://www.bpnews.net/10832/opponents-of-pornography-contend-its-not-harmless-or-victimless

Pizarro, D. A., Inbar, Y., & Helion, C. (2011). On disgust and moral judgment. *Emotion Review*, *3*(3), 267 –268. http://doi.org/10.1177/1754073911402394

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57–75. http://doi.org/10.1037/a0021867

Ritter, R. S., & Preston, J. L. (2011). Gross gods and icky atheism: Disgust responses to rejected religious beliefs. *Journal of Experimental Social Psychology*, *47*(6), 1225–1230. http://doi.org/10.1016/j.jesp.2011.05.006

Rottman, J., Kelemen, D., & Young, L. (2014). Tainting the soul: Purity concerns predict moral judgments of suicide. *Cognition*, *130*(2), 217–226. http://doi.org/10.1016/j.cognition.2013.11.007

Rottman, J., & Young, L. (2014). Comment: Scholarly disgust and related mysteries. *Emotion Review*, *6*(3), 222–223.

Royzman, E., Atanasov, P., Landy, J. F., Parks, A., & Gepty, A. (2014). CAD or MAD? Anger (not disgust) as the predominant response to pathogen-free violations of the divinity code. *Emotion*, *14*(5), 892–907. http://doi.org/10.1037/a0036829

Royzman, E., Goodwin, G. P., & Leeman, R. F. (2011). When sentimental rules collide:"Norms with feelings" in the dilemmatic context. *Cognition*, *121*(1), 101–114.

Royzman, E., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, *10*(4), 296–313.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*(4), 574–586. http://doi.org/10.1037/0022-3514.76.4.574

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145.

Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: what it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological Bulletin*, *139*(2), 328–351. http://doi.org/10.1037/a0029319

Schein, C., Goranson, A., & Gray, K. (2015). The uncensored truth about morality. *The Psychologist*, *28*(12), 982–985.

Schein, C., & Gray, K. (in press). Moralization and harmification: The dyadic loop explains how the innocuous becomes harmful and wrong. *Psychological Inquiry*.

Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, *41*(8), 1147–1163. http://doi.org/10.1177/0146167215591501

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, *34*(8), 1096–1109. http://doi.org/10.1177/0146167208317771

Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought-action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders*, *10*(5), 379–391. http://doi.org/10.1016/0887-6185(96)00018-7

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.

Sood, A. M., & Darley, J. M. (2007). The plasticity of harm in the service of punishment goals: An experimental demonstration. Presented at the 5th Annual Conference on Empirical Legal Studies Paper, Yale Law School: Wiley-Blackwell.

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

Terrizzi Jr, J. A., Shook, N. J., & Ventis, W. L. (2010). Disgust: A predictor of social conservatism and prejudicial attitudes toward homosexuals. *Personality and Individual Differences*, *49*(6), 587–592.

Thórisdóttir, H., & Jost, J. T. (2011). Motivated Closed-Mindedness Mediates the Effect of Threat on Political Conservatism. *Political Psychology*, *32*(5), 785–811. http://doi.org/10.1111/j.1467-9221.2011.00840.x

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

Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology*, *97*(1), 103.

Tybur, J. M., Merriman, L. A., Hooper, A. E., McDonald, M. M., & Navarrete, C. D. (2009). Extending the the behavioral immune system to political psychology: are political

conservatism and disgust sensitivity really related? *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, *8*(4), 599–616.

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

Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, *17*(6), 476–477. http://doi.org/10.1111/j.1467-9280.2006.01731.x

van Leeuwen, F., & Park, J. H. (2009). Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences*, *47*(3), 169–173. http://doi.org/10.1016/j.paid.2009.02.017

van Leeuwen, F., Park, J. H., Koenig, B. L., & Graham, J. (2012). Regional variation in pathogen prevalence predicts endorsement of group-focused moral concerns. *Evolution and Human Behavior*, *33*(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.

Wheelerr, C. (2015, August 23). Children were cheap guinea pigs for deadly tainted blood products | UK | News | Daily Express. Retrieved from http://www.express.co.uk/news/uk/600114/Children-were-cheap-guinea-pigs-deadly-tainted-blood-products

Wright, J. C., & Baril, G. L. (2013). Understanding the role of dispositional and situational threat sensitivity in our moral judgments. *Journal of Moral Education*, *42*(3), 383–397. http://doi.org/10.1080/03057240.2013.820659

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

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using

hierarchical linear models problems and solutions. *Organizational Research Methods*,

12(4), 695–719.

Figures

Figure 1. Different models that link disgust, harm and immorality. The historically popular

direct disgust model (Haidt & Hersh, 2001) suggests that perceptions of harm are epiphenomenal

and unimportant in the moral condemnation of disgusting acts. Conversely, *dyadic morality*

(Schein & Gray, 2015) suggests that perceptions of harm are the most important "active

ingredient" in moral judgment. That is, harm should mediate the relationship between disgust

and moral condemnation, even for disgusting-but-ostensibly-harmless acts. (Solid lines represent

hypothesized causal links, dotted lines represent non-causal links).
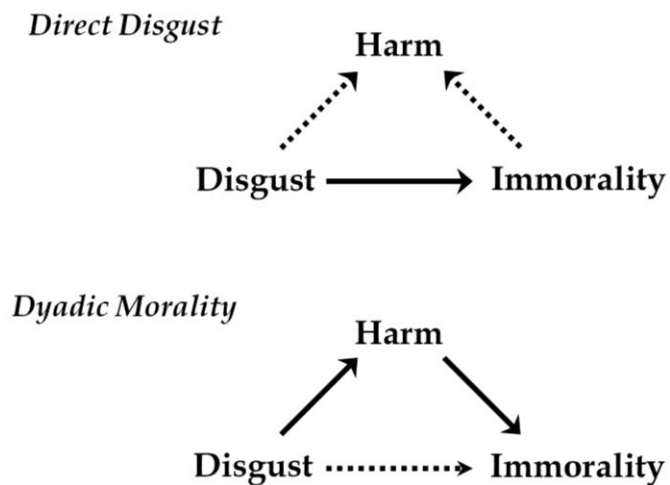
Figure 2. Mediation of the effect of disgust sensitivity on explicit disapproval of same-sex marriage through perceived harm (Study 1). Model controls for politics and age. * p < .001
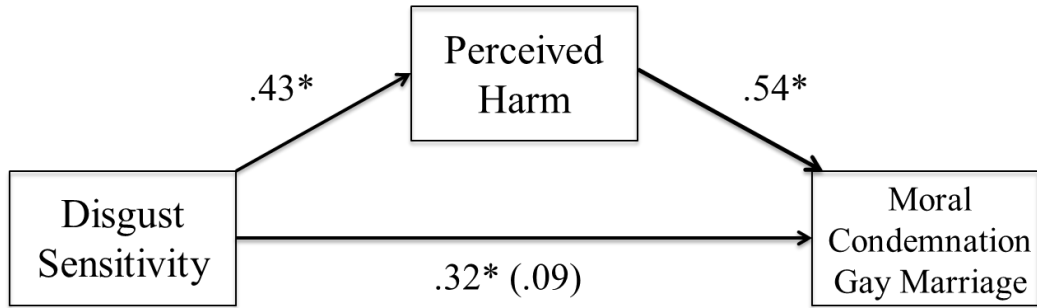
Figure 3. Mediation of the effect of disgust sensitivity on implicit disapproval of gay intimacy

through perceived harm, controlling for politics and age (Study 1). * $p < .001$
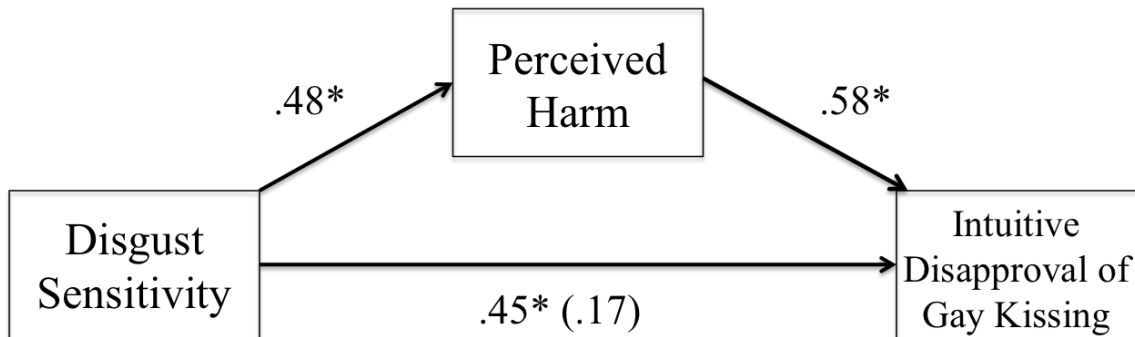
Figure 4. Mediation of the effect of disgust on moral judgment through perceived harm controlling for Christian identity (Study 2). *p < .001.
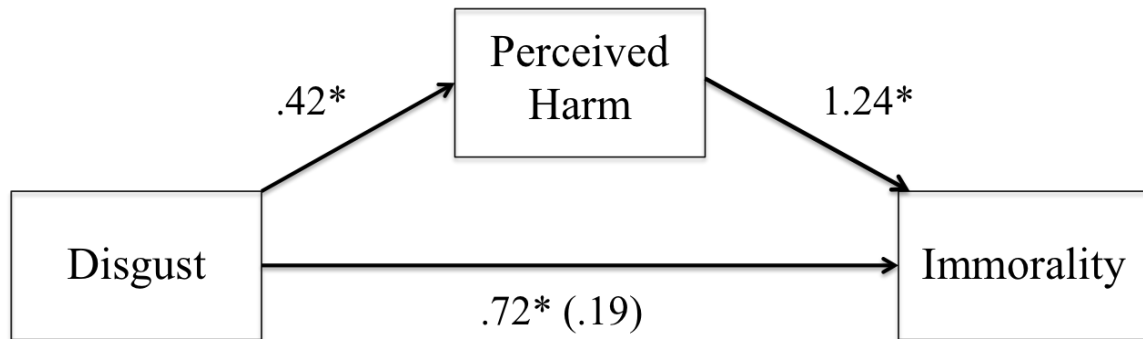
Figure 5. MLM fixed effect coefficients for the relationship between disgust and immorality
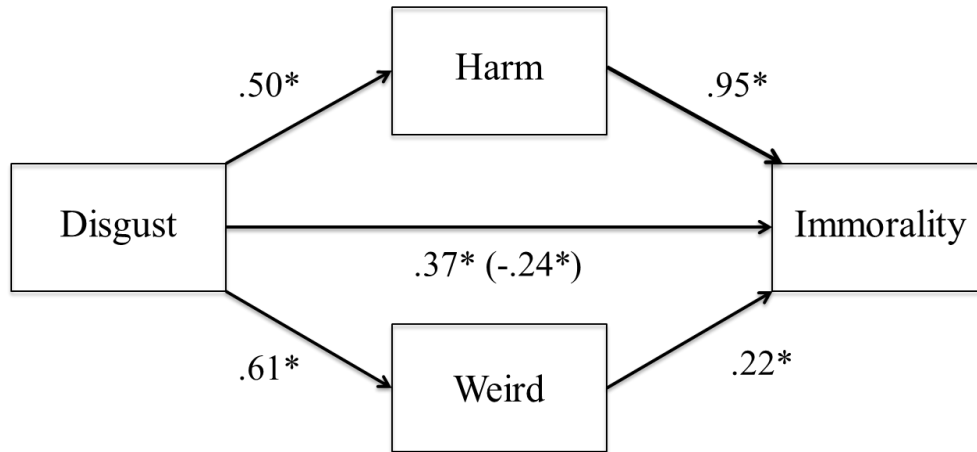
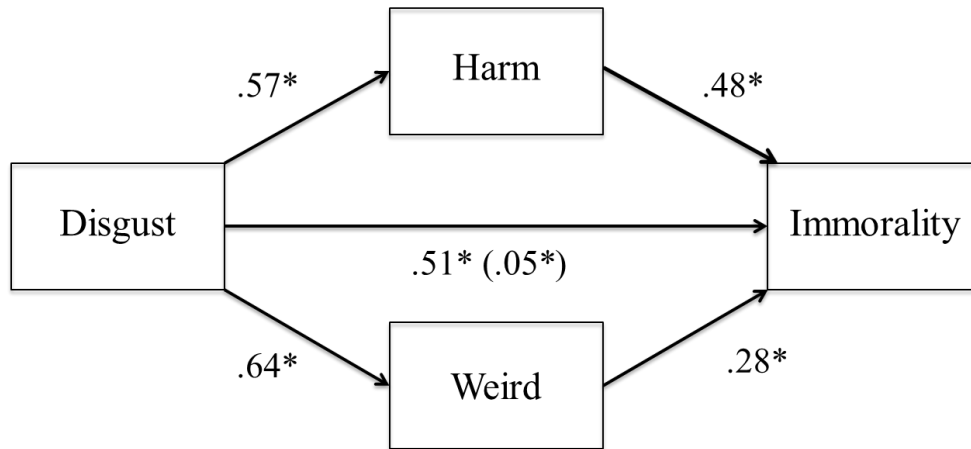mediated by harm and weirdness (Study 3). * p < .001.

Figure 6. MLM fixed effect coefficients for the relationship between perceived disgust and immorality mediated by perceived harm and weirdness excluding the canonically harmful immoral acts (Study 3). * *p* < .01

Footnotes

[1] perhaps except funny gross-out scenes in comedy movies

[2] Haidt and Hersh (2001) claims that disgust "drives" the moral condemnation of sexual practices but uses only correlational analyses.

[3] The MFQ Relevance scale goes even further than mere self-reports, asking participants to introspect upon the bases of their own moral judgment, which decades of work in social psychology suggests is unreliable (Nisbett & Wilson, 1977)