

Is Purity a Distinct and Homogeneous Domain in Moral Psychology?

Dolichan Kollareth¹, Hiram Brownell¹, Juan Ignacio Durán², & James A. Russell¹

¹Boston College

²Universidad a Distancia de Madrid

© 2022, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI: 10.1037/xge0001274

Author Note

Dolichan Kollareth, Hiram Brownell, James A. Russell, Department of Psychology and Neuroscience, Boston College. Juan Ignacio Durán, Universidad a Distancia de Madrid

For data from each of the studies, see

https://osf.io/2rm47/?view_only=b221b57882f744029b938d61c9e93ec1

Correspondence concerning this article should be addressed to Dolichan Kollareth, Department of Psychology and Neuroscience, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA. Email: dollychank@gmail.com

Abstract

“No” is our answer to the question in our title. In moral psychology, a purity violation (defined as an immoral act committed against one’s own body or soul) was theorized to be a homogeneous moral domain qualitatively distinct from other moral domains. In contrast, we hypothesized heterogeneity rather than homogeneity, overlapping rather than distinct domains, and quantitative rather than qualitative differences from other hypothesized domains (specifically, autonomy, which is harm to others). Purity has been said to consist of norms violations of which elicit disgust and taint the soul. Here we empirically examined homogeneity: whether violations of body (e.g., eating putrid food) belong in the same moral domain as violations of the soul unrelated to bodily health (e.g. selling one’s soul, desecrating sacred books). We examined distinctness: whether reactions to purity violations differ in predicted ways from those to violations of autonomy. In four studies (the last pre-registered), American internet users (in Studies 2 and 4, classified as politically conservative or liberal) ($N_s = 80, 96, 1312, 376$) were given stories about violations based on prior studies. Non-health purity violations were rated as relatively more *disgusting*, but less *gross* (the lay term for the reaction to putrid things), and more likely to taint the soul than were health-related ones. Surprisingly, both health and non-health purity violations were typically judged as only slightly immoral if at all. Autonomy violations were rated as more *disgusting* and tainting of the soul than were purity violations.

Keywords: morality-immorality judgment; emotion; purity; disgust; autonomy

Is Purity a Distinct and Homogeneous Domain
in Moral Psychology?

A far-reaching theoretical development in moral psychology is the division of morality into a small number of qualitatively distinct domains (Rozin et al., 1999). The focus of the present study is one such domain -- called *divinity*, *sacred*, or *purity*. We question whether purity qualifies as a distinct and homogeneous domain.

The present study is part of a larger concern with the general approach to descriptive morality of postulating fundamentally distinct domains (Graham et al., 2013; Haidt & Joseph, 2007). There are at least two alternatives to positing distinct domains. In one, all morality has a single basis; candidates include the Golden rule, Kant's categorical imperative, Bentham's utilitarianism, and the recent proposal that harm is the common denominator (Schein & Gray, 2015; 2018).

Another alternative is that morality consists of a very large number of overlapping and crisscrossing rules for behavior, rules with no common features shared by all or even by distinct clusters but that instead vary with historical period, culture, context, and social role (see Bucciarelli et al., 2008; McHugh et al., 2022, for such a view). Thus, rather than postulating a distinct moral domain of purity, a constructionist view of the human mind might provide an alternative explanation of judgement differences between certain unusual scenarios like sex with a corpse and other moral scenarios (Gray et al., 2022; Cameron et al., 2015). That is, moral judgments are subserved by domain-general processes: similar psychological mechanisms underlie moral judgement of purity and other moral violations (Gray et al., 2022; Cameron et al., 2015).

Such broad issues are fundamental to the empirical study of morality, but here we begin with one preliminary question raised by Gray et al. (2021) within this broader set of issues: whether purity is a homogeneous domain distinct from autonomy, another alleged moral domain. We suggest that purity is heterogeneous and that purity and autonomy might be better understood not as qualitatively distinct but as variations along a continuum.

The Assumptions of Homogeneity and Distinctness

To our knowledge, the concept of moral domain in general is not precisely defined. In the realm of descriptive morality, the necessary and sufficient properties that distinguish one domain from another are not consensually agreed upon, but implicitly the notion is that violations within a domain are homogeneous – that is, they all share common properties -- and that these properties are qualitatively distinct from those in any other domain.

Moral Emotions. Different theorists characterize the separate domains somewhat differently, but one seminal idea is that each domain is characterized by a distinct “basic” emotion elicited in a witness to the violation (Rozin et al., 1999). Later theories continue to highlight the idea that each domain is linked to a different “moral emotion” (Haidt & Joseph, 2007 p. 380; Young & Tsoi, 2013 p. 591; Li et al., 2016 p. 26) or an “automatic affective evaluation” (Graham et al., 2013 p. 109). A common assumption is that purity violations elicit disgust (a reaction that in lay terms is feeling grossed-out), whereas autonomy violations (harm to another person) elicit anger. The witness to the purity violation, in turn, rationalizes the disgust reaction as signaling a violation of a sacred moral code. Sometimes, disgust is implied to play a causal role in the subsequent moral judgment (Seidel & Prinz, 2013).

On this theory, disgust is a basic emotion. It is a biologically based reaction in which an ancient rejection of threats of pathogens, parasites, and toxins that pose a danger to physical

health was co-opted over the course of evolution to reject threats to purity unrelated to physical health—threats to spiritual purity or that taint one’s soul or that degrade the sacred elements of one’s self (Graham et al., 2013; Haidt & Graham, 2007; Horberg et al., 2009; Milesi, 2016; Rozin et al., 2008) – as seen in other names for this domain: *divinity* and *sacred*. We question whether, in fact, threats to one’s physical health and threats to one’s non-physical purity (i.e., sacred values) both elicit the same emotion.

Taint to Soul. Purity violations have been said to degrade one’s spiritual self (Graham et al., 2013; Horberg et al., 2009). Consistent with this idea, Rottman et al. (2014) offered “taint to soul” as a marker of a purity violation. Although not offered as a defining feature, it has been hypothesized that those with a conservative political orientation, perhaps because of enhanced concerns about spiritual self, are more condemning of purity violations than are those with a more liberal orientation (Graham et al., 2009). We question whether, in fact, purity violations are the only type of moral violation that a witness believes taints or degrades the soul.

Homogeneity: Physical Health and Non-health. The actions used in research as examples of purity violations have been diverse: touching dog feces, selling one’s soul to the devil, desecrating religious symbols, cleaning the bathroom with the national flag, masturbating with a dead chicken, reading scripture while defecating, and wearing mismatched clothing (Cannon et al., 2011; Dungan et al., 2017; Guerra & Giner-Sorolla, 2010; Horberg et al., 2009; Vasquez et al., 2001). Some of these actions, such as touching dog feces, threaten physical health (we call these actions physical health-related). Other actions, such as selling one’s soul to the devil or desecrating religious symbols, however, are not related to physical health. They are more heterogeneous and often associated with soul, spirit, and the like: Theorists have characterized this duality in various ways by describing purity norms as a guardian of the “body

as well as soul” (Rozin et al., 2008), “body and spirit” (Cannon et al., 2011), “physical and mental” (Horberg et al., 2011), “bodily...and religious” (Haidt & Graham, 2007), “physical and spiritual” (Haidt et al., 2009; Koleva et al., 2014; Simpson & Rios, 2017), “biological and social” (Federico et al., 2016). No single term is used for the non-health purity norms. And, we do not find a clear and consensual definition for those actions that constitute purity violations unrelated to physical health. We shall simply call them non-health purity violations. Whatever the name, the homogeneity hypothesis to be tested is that health-related and non-health-related purity violations share the same predicted properties. Our alternative hypothesis is that threats to one’s physical health and threats to one’s non-physical purity differ in terms of witnesses’ emotional reaction (specifically in feeling grossed-out) and their perception of taint to the soul of the perpetrator.

We suggest that the purity domain contains at least two separate sets of norms for behavior, probably more, violations of which do not elicit precisely the same emotional profile. One set of norms stems from biological imperatives, principally the behavioral immune system, but possibly others. The behavioral immune system has strong empirical support (Murry & Schaller, 2016; Schaller, 2016; Schaller & Park, 2011). It evolved with the function of avoiding disease, parasites, and toxins, not just in humans but, in some form or other, in other species as well; for example, rats avoid a sick rat (Rieger et al., 2022).

The other set of norms, for non-health purity violations, apparently stems largely from ideology, often religious or quasi-religious or political. Examples used include selling one’s soul, worshipping the devil, violating religious instructions, desecrating sacred objects, endorsing Adolf Hitler or a child molester, behaving like an animal, or tampering with the natural order of things by, for example, implanting horns in one’s skull. Consistent with non-health purity norms

stemming from ideology is evidence that conservatives are more condemning of purity violations than are liberals (Graham et al., 2009). Health and non-health norms can and sometimes do overlap, but sometimes can and do conflict. In proposing differences within the purity domain, we do not mean to be taken as advocating two distinct domains in place of one; indeed, we doubt that dividing morality into distinct domains is a useful theoretical step.

Distinctness. Much research continues to assume that purity is a distinct moral domain. Some of this research is basic science (Anderson & Masicampo, 2017; Berniūnas et al., 2021; Chakroff et al., 2017; Dungan et al., 2017; Dungan & Young, 2019; Haidt, 2012; Heerdink et al., 2019; Liuzza et al., 2019; Ritter et al., 2016; Rottman & Young, 2019; Sweetman & Newman, 2020; Tracy et al., 2019). Other of this research is applied science (Arden & Rabinovitz, 2020; Berman & Small, 2018; Kennedy et al., 2021; Koleva et al., 2012; Pilotti & El Alaoui, 2019; Rottman et al., 2014; Scott et al., 2016; Vezzali et al., 2017).

Nonetheless, other studies of purity question its distinctness. For example, participants reacted with more anger than disgust to some purity violations – just as they do to autonomy violations (Franchin et al., 2019). Participants did not always judge religious purity violations free from threats to health as disgusting (Royzman et al., 2014; Kollareth & Russell, 2019). Actions depicted as purity violations can be perceived as harming others and thus may not be distinct from autonomy violations (Gray, 2014; Schein & Gray, 2015; Schein et al., 2016).

Evaluating the Evidence

Conceptual Issues. Evaluating any evidence challenging the purity domain is difficult because neither “domain” nor “purity” is well defined (and, indeed, theorists disagree on the purity domain’s properties; Chakroff et al., 2013). Lack of agreed upon definition hinders empirical study of the theory because researchers cannot know definitely which events to study:

which are and which are not violations of purity to include in research. Often the religious violations used in studies that challenge purity's homogeneity were newly created rather than taken from previous literature that ostensibly provided support for the purity domain (Royzman et al., 2014; Kollareth & Russell, 2019). Thus, in the absence of a clear definition of a purity violation, one could question whether the events studied are regarded as genuine violations of the purity domain by those who posit that domain. By contrast, in the present investigation, we relied on stories or norms from previous research advocating purity as a separate domain.

Confounds. In reviewing previous studies on the purity domain, we found three methodological issues problematic, and we here offer what we consider to be improvements. First, some purity violations confound a physical health-related element with a non-health violation of a religious, political, social, or sacred norm: for example, thinking of sacred scripture while defecating. This event was judged as both disgusting and immoral (Haidt, 2003), but it's possible that it is the excrement that elicits disgust, but it is the thinking of scripture in an inappropriate context that violates a sacred norm. Another type of violation examined was sexual (incest, necrophilia) in which a potential danger to health and other biological consequences is combined with a religious/social norm violation concerning appropriate sexual behavior. In yet another type of violation, the action depicted is so unusual as to be nonconformist and abnormal (masturbation with a dead chicken). In some studies, responses to violations involving physical health were averaged with non-health purity violations, allowing no conclusion about whether responses to these two types of violations were similar or different (see Cannon et al., 2011; Dungan et al., 2017; Horberg et al., 2009). The question whether health and non-health violations belong in the same purity domain requires de-confounding these types of actions and

then examining responses to each type separately. Here we did so, at least to a preliminary degree.

The term *Disgust*. Second, studies of the purity domain typically rely on the term *disgust* in the assessment of the emotion theorized to be commonly elicited by purity violations and only purity violations, but recent findings showed that the term *disgust* is used by lay judges not only for unhealthy things but in other contexts more generally for strong negative emotional reactions (Marzillier & Davey, 2004; Hutcherson & Gross, 2011; Royzman & Kurzban, 2011; Rozin et al., 2009). A dictionary included *disapproval* and *dislike* as part of the meaning of *disgust* (Landau, 2000). Instead, the word *grossed-out* may be used by lay persons to refer to a genuine core disgust reaction to unhealthy things (Herz & Hinds, 2013; Kollareth & Russell, 2017; Nabi, 2002). A convincing demonstration that people react with the same emotion to a moral violation as they do to a simply putrid event should examine ratings of *grossed-out* as well as of *disgust*. We tested the plausible hypothesis that health-related purity violations elicit principally a feeling of being *grossed-out* in the witness, but that non-health purity violations (often offenses against religious or sacred norms) elicit principally a feeling of *disgust* – in the sense of disapproval or outrage—in the witness. Further, we anticipated that if *disgust* and *grossed-out* have the meanings we suggest, autonomy violation will also be rated as *disgusting* but not as *gross*.

Are the violations immoral? Third, surprisingly, prior studies did not always provide a convincing demonstration that the alleged purity violations included were perceived as immoral. See Supplement 1 for some previously used immorality rating scales. Some studies of alleged purity violations had no measure of immorality (Guerra & Giner-Sorolla, 2010; Vasquez et al., 2001). Some studies used a measure, not of immorality, but of negativity - positivity (Cannon et

al., 2011). Although negativity-positivity is important, it is not the same as morality. Catching the flu is negative, but not immoral. In one of their studies, Piazza et al. (2019) assessed right and wrong. Believing that the moon is made of green cheese would be wrong, but not immoral. Still other studies did assess immorality (Dungan et al., 2017; Horberg et al., 2009), but did so with a *unipolar* response format. That is, the rating scale ostensibly assessed degrees of immorality, but not morality. Two problems exist in any unipolar response format when used for a variable with a clear bipolar opposite (moral). Because of random noise in any measurement, the response distribution to the unipolar scale will almost inevitably include scores above the “not immoral” anchor, seemingly showing some degree of immorality even if no participant truly found the act to be immoral. Second, different participants interpret the typical unipolar format differently; they place the neutral point (neither moral nor immoral) at different locations along the scale, with some placing it toward the middle (Russell & Carroll, 1999). When the neutral point is implicitly placed away from the bottom anchor, ratings below that neutral point indicate morality rather than immorality. If so, the researchers cannot conclude that the violation was truly judged immoral. So, in the present study, we used a bipolar response scale specifically for morality-immorality. The neutral point is specified explicitly. (For dimensions that require a unipolar scale, such as that for a categorical emotion, we use the “explicitly unipolar” scales suggested by Russell and Carroll, 1999).

Overview

In four studies, we critically examined the hypothesis that purity is a homogeneous and distinct moral domain. In contrast, we specifically hypothesized that health and non-health violations elicit different patterns of reactions in a witness. In three of the four studies, we critically examined the hypothesis that purity is distinct from the allegedly distinct moral domain

of autonomy. Thus, overall, we examined three types of alleged moral violations: health-related purity, non-health purity, and autonomy violations – all based on stories used in previous research to exemplify these types of violations. We also re-examined the role of conservative as opposed to liberal orientation in these moral judgments. After reading a short story about each violation, the participants rated how each violation made them feel on scales of *disgust*, *grossed-out*, and *anger* (plus emotion comparison items). Participants also rated how much the violation was moral or immoral. In Studies 1, 3, and 4, participants rated how much each act tainted the soul of the protagonist. In Studies 2 and 4, political orientation was assessed.

The studies reported here were approved by the Boston College Institutional Review Board: Emotional Reactions to Moral Violations (protocol number: 18.047.01). Participants in the studies provided consent online. An online supplementary document (Supplementary Procedures) provides the stories, instructions, and the scales used in each of the studies. For data from each of the studies, see

https://osf.io/2rm47/?view_only=b221b57882f744029b938d61c9e93ec1 (Kollareth, 2021).

Study 1: Health and Non-health

Purity Violations

Here we examine actions designated in prior studies as purity violations, but selected here to be clearly either physical-health-related or non-health-related. Half of the participants read about a protagonist named Frank and ten actions that he committed. Three stories were filler, and seven were taken from previous research to depict purity violations: three of the purity violations were threats to physical health, and four were not related to health. The other half of the participants followed a similar procedure by reading about a protagonist named Marianne and her ten actions (all but one different from those carried out by Frank). This second half of the

study was thus a systematic replication of the first half but with a female protagonist and different acts. To minimize confounding of purity violations with violations that harm others, both protagonists were described as living alone.

The filler stories served multiple purposes. Because they were identical in format to the experimental story, they familiarized the participant with the procedure. Because some filler stories described a moral, and others an immoral, deed, they set a broader context against which the experimental story would be judged. Stories describing health and non-health violations in both sets were taken from previous studies. We examined four previous publications by different researchers who offered examples of violations of purity norms (including sacred, political, and social symbols) (Cannon et al., 2011; Dungan et al., 2017; Guerra & Giner-Sorolla, 2010; Horberg et al., 2009). We excluded examples (a) in which health and non-health were confounded, (b) in which the action was highly unusual (and thus highly non-conformist, possibly violating a community moral norm), or (c) in which the action could not be described in a simple story.

Participants rated how each act made them feel on scales of *disgust*, *grossed-out*, *anger* and also *sadness* and *irritation*. As a measure of moral purity specifically, we used a rating scale from Rottman et al. (2014), who asked their participants whether committing the action tainted the person's soul. Participants also rated how much the act was moral or immoral along a single bipolar dimension.

Method

Participants. Participants ($N = 80$) were Americans and native English speakers recruited through Amazon Mechanical Turk (50 women, 30 men; $M_{\text{age}} = 37.81$ years, range: 22-70 years). Our power analysis showed that for the series of planned contrasts (such as *anger* vs.

disgust), 80 participants were needed for a medium sized effect ($d = .5$) with the conventional power value of $(1-\beta) = .80$, with a two-tailed alpha of $.05$. All participants passed an attention check by providing a unique code at the end of the survey and by typing the name of their favorite color; and so all data collected are reported here.

Stories. There were two questionnaires. In one, the protagonist was Frank, who lived alone on a very small retirement income. In the other, the protagonist was Marianne, who lived alone on a small farm she inherited. In the set with Frank as protagonist, three stories described health purity violations and four stories described non-health purity violations. In the set with Marianne as protagonist, four stories described health purity violations and three stories described non-health purity violations. There were also three filler stories in each set. Supplementary Procedures 1 provides the two sets of questionnaires and the publication from which the health purity and non-health purity stories were taken.

Due to experimenter error, one story (eating in the bathroom) was repeated in both the Frank and the Marianne datasets. Data from the repeated story were deleted from the Marianne data set prior to the analyses reported below. Analyses were thus based on 7 unique stories (3 health, 4 non-health) in the Frank dataset and 6 unique stories (3 health, 3 non-health) in the Marianne dataset. See Table 1 for the health and non-health stories.

Table 1

Health and non-health purity scenarios, Study 1

| Health scenarios |
|---|
| Frank cooked and ate his dog after it died of natural causes |
| Frank ate an unwrapped chocolate bar that he found in the dustbin |

Frank ate in the same place that he goes to the bathroom

Marianne licked between her toes

Marianne picked up dog poop in the barn with her bare hands

Marianne stepped in vomit barefoot

Non-health scenarios

Frank signed a piece of paper selling his soul on a dare

Frank wrote “666” (the number of the beast, a reference to the devil) in his hymn book and bible

Frank spent a day wearing a sweater once owned by Adolf Hitler

Frank wore unmatched clothing

Marianne opposed the religious authority of her group by refusing to sign a letter to be sent to a newspaper

Marianne ate certain foods even though they are forbidden in the holy scriptures of her group

Marianne uses pesticides in her farming even though her religious group says that pesticides are unnatural and against nature’s laws and therefore forbidden

Unipolar emotion response scales. The emotion response format had 5 emotion labels: *anger, disgust, sadness, grossed-out, irritation*. Participants could choose as many or as few labels as they wanted by circling “yes” or “no” for each. For any emotion label that was circled “yes,” the participant was asked to rate its intensity with a 7-point Likert scale (ranging from 1-*barely* to 7-*extremely*). The score for each emotion label was 0 for “no” and the intensity rating for “yes.” Thus, the range was 0 to 7.

Unipolar scale for tainting the soul. A scale for tainting the soul—a measure of purity—was taken from Rottman et al. (2014). Participants were asked: “Was the purity of [Frank’s / Marianne’s] soul tainted as a result of [his/her] act? If yes, how much?” Participants could choose by clicking “yes” or “no” for the question. If chosen by clicking “yes,” the participant was asked to rate its intensity with a 7-point Likert scale (ranging from 1-*slightly* to 7-*highly*). The score for ‘tainting the soul’ was 0 for “no” and the intensity rating for “yes.” Thus, the range was 0 to 7.

Bipolar scale for morality-immorality. Participants were asked: “How moral or immoral is this act?” The response format was a 7-point bipolar scale (1-*extremely moral*, 2-*moderately moral*, 3-*barely moral*, 4-*neutral*, 5-*barely immoral*, 6-*moderately immoral*, and 7-*extremely immoral*).

Design. The factor of greatest theoretical interest—the distinction between health and non-health violations—was within subjects. Less relevant factors such as the gender of the protagonist in the stories and the gender of the participant were between-subjects.

Procedure. Participants completed the survey online. They were randomly assigned to the questionnaire with Frank or with Marianne as protagonist. The participant first read the consent form, then the general instructions, and then ten stories. Participants read the three filler stories first. The order of the other seven stories and the order of the emotion labels in the emotion response scale were randomized separately for each participant. So, each participant received a different order of non-filler stories and a different order of emotion labels. After reading of each story, the participant completed the emotion response scales, tainting the soul scale, and then the immorality scale, in that order.

Results and Discussion

Check on the bipolar morality-immorality rating scale. Participant responses to the filler stories showed that the bipolar morality-immorality rating scale worked well. Participants judged the moral filler story as *moral* ($M = 1.68$, $SD = 0.96$, between *moderately* and *extremely*), and they judged the immoral filler story as *immoral* ($M = 5.05$, $SD = 1.33$, close to *barely*), but the neutral story as between *barely* and *moderately moral* ($M = 2.95$, $SD = 1.43$). (See Supplement S1-1 for ratings for each of the measures for each of the stories).

Emotional responses. The theory of a homogeneous purity domain predicts that purity violations – whether health or non-health -- are disgusting (with *gross* the best descriptor of that feeling by lay persons) rather than angering. In contrast, we predicted distinct emotional responses for health and non-health purity violations. Figure 1 presents ratings for each of the three most relevant emotion terms (*disgust*, *grossed-out*, *anger*) separately for health and non-health violations. (Results from the other emotions are presented in Supplement S1-2). Visually, health and non-health elicited distinct emotional responses; this appearance was confirmed by a violation type (2) x emotion (3) Analysis of Variance (ANOVA), $F(2, 158) = 175.72$, $p < .001$, partial eta squared .69. Health purity violations were characterized by high levels of both *grossed-out* and *disgust*, and very little *anger*, whereas non-health purity violations were characterized by negligible degrees of *disgust*, *anger*, and, especially, *grossed-out*.

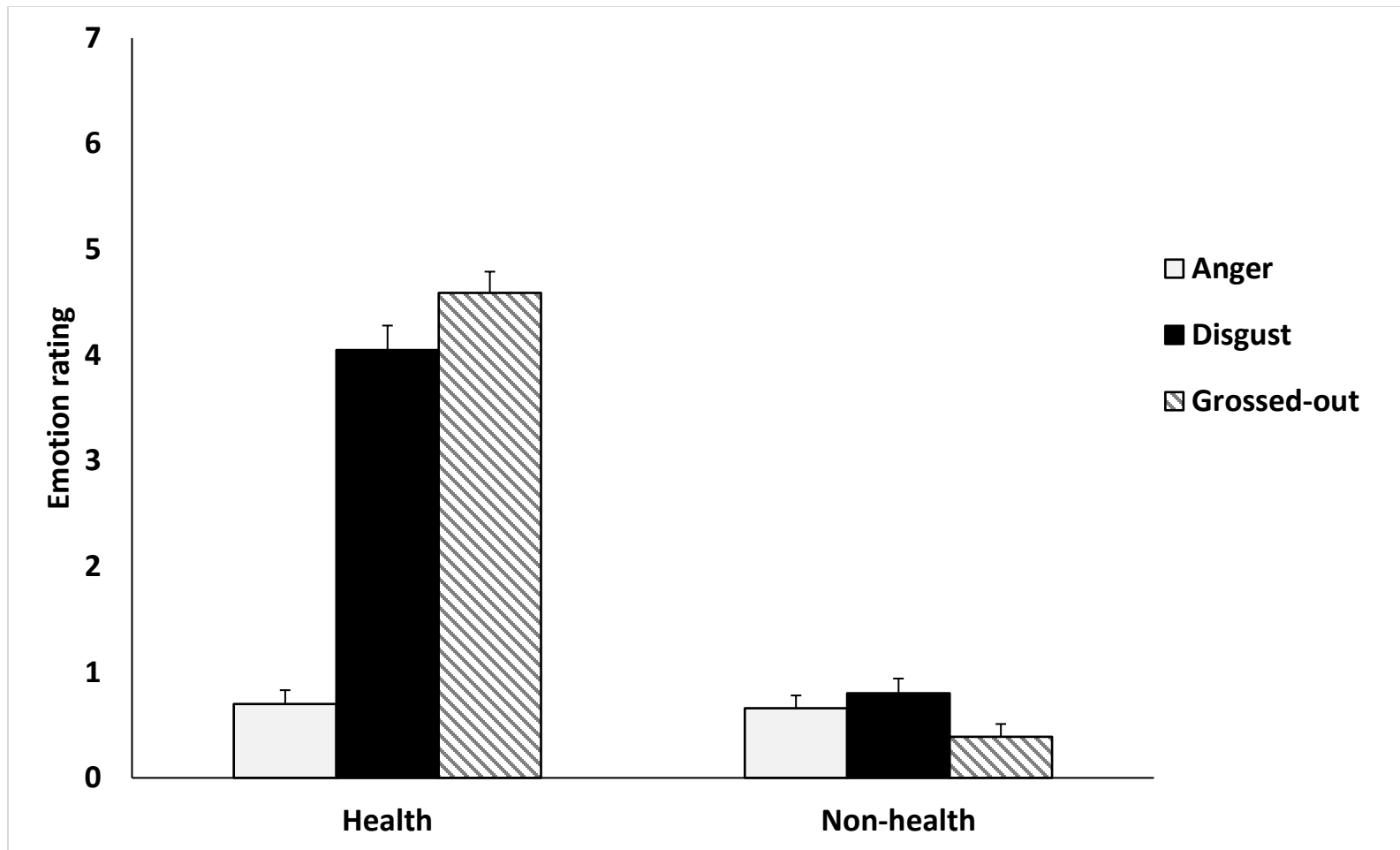


Figure 1. Mean emotion ratings for health and non-health purity violations, Study 1. Each emotion was rated on an 8-point scale with 0 defined as “no” and 7 as extreme. Standard errors are represented in the figure by the error bars attached to each column.

Further (2 x 2) ANOVAs, each with the two purity violation types (health, non-health) and two levels of emotion (*anger* vs. *disgust* in one, and *disgust* vs. *grossed-out* in the other), confirmed the reliability of the interaction pattern shown in Figure 1 by revealing reliable interactions for *anger* versus *disgust*, $F(1, 79) = 158.03, p < .001$, partial eta squared = .67, and for *disgust* versus *grossed-out*, $F(1, 79) = 23.76, p < .001$, partial eta squared = .23. This last interaction show that health-related purity violations elicit principally a feeling of being *grossed-out* in the witness, but that non-health purity violations elicit little emotion, but with *disgust* the highest. For non-health stories, due to a floor effect in the data, we also conducted a sign test, which supported the same conclusion: *disgust* was higher than *grossed-out* for 23 out of 25 participants with 55 ties excluded, $z = 4.20, p < .001$. Results for emotion ratings within each story separately for health and non-health violations are given in Supplement S1-3. These results justify generalizability of our findings.

In sum, *disgust*, *grossed-out*, and *anger* did not show an identical pattern across health and non-health purity violations. Participants reacted with more *disgust* and *grossed-out* but with little *anger* for health purity violations, whereas they reacted with overall little emotion, least of all *grossed-out*, for non-health purity violations. Distinguishing *disgust* from *grossed-out* also showed that participants respond differently to the two types of purity violations. Participants used *disgust* and *grossed-out* in a way that dissociated health from non-health purity violations. Indeed, there was no convincing support for the hypothesis that witnesses found the non-health purity violations *gross* to any appreciable degree (i.e., more than other emotions assessed, even positive ones) even though *grossed-out* appears to come closest to expressing the core emotion Rozin et al. (1999) theorized as the basis for moral condemnation of purity violations.

Tainting of the soul. The theory of a purity domain predicts that violations taint the soul. Our participants found non-health purity violations to be more tainting to the soul than are health purity violations. In a dependent sample t-test, non-health purity violations ($M = 1.11$, $SD = 1.44$) were significantly more tainting to soul than health-purity violations ($M = 0.59$, $SD = 0.92$), $t(79) = 3.99$, $p < .001$, partial eta squared = .40.

Due to the low ratings for tainting of the soul, we also report an alternative analysis based on the proportion of participants who found a violation tainting the soul. Of the 80 participants, 36 denied any taint whatsoever to any purity violation, either health or non-health, and only 22 assigned at least some taint to both violation types. Twenty five participants found some taint from health violations, and 41 participants found some taint from non-health violations. Taint was more consistently attributed to non-health than to health violations: By a sign test, taint to soul attribution was higher for non-health violations than for health violations for 35 out of 44 participants (with 36 ties excluded), $z = 3.77$, $p < .001$. Thus, purity violations were not always seen as tainting the soul. Non-health violations were more tainting than health violations, which were rarely seen as tainting.

A multilevel mixed-effect regression model was used to predict taint to soul ratings as the outcome variable distinguishing between variations across two nested levels: participants and stories. The model was designed to predict variability between participants in outcome from immorality ratings while controlling for the effect of the type of purity violation (health vs. non-health) in outcome between stories. Quantitative variables were standardized before performing the analysis to obtain regression weights (β) that could be interpreted as partial correlations. According to the model, tainting to soul was significantly correlated with immorality ($\beta = .41$, $t(12.35) = 3.07$, $p = .009$).

Morality-immorality. The theory of a purity domain assumes that violations – whether health-related or not – are immoral. On our measure of immorality, predictions of the standard theory of a purity domain received ambiguous support. On the one hand, participants judged both physical health and non-health purity violations to be similar on *morality-immorality*. In a paired sample t-test, there was no significant difference in the mean ratings for immorality of health ($M = 4.42$, $SD = 0.91$) versus non-health ($M = 4.46$, $SD = 0.86$), $t(79) = 0.48$, $p = .640$. On the other hand, the mean immorality rating for both types of violation fell between *neutral* (scale value of 4) and *barely immoral* (scale value of 5). Mean ratings on the *morality-immorality* scale for each of the stories are provided in Supplement S1-1. Participants found 9 of the 13 non-redundant purity violations taken from previous research less immoral than “barely.” Indeed, three of these nine violations were on the moral side (<4) of *neutral*. Only four of the 13 were judged as or more immoral than “barely” and of those only one was significantly so (Frank cooked and ate his dog after it died of natural causes).

Other correlational analyses. On some analyses (Rottman et al. 2014), the strength of the individual’s emotional reaction of disgust (but of no other emotion) is the predicted basis for a judgment of immorality for purity violations – whether health or non-health related. That is, the greater the disgust the more immoral.

This hypothesis was tested with a mixed-effect multilevel regression. Both quantitative and dichotomous variables can be included as predictors in a regression model. The tests corresponding to quantitative predictors can be interpreted in terms of the existence of a linear relationship between the predictor and outcome variables and the tests on dichotomous predictors can be interpreted in terms of the existence of a mean difference in the outcome between the categories specified by the predictor variable. Finally, tests on interaction effects between

quantitative and dichotomous predictors can be interpreted as the mean difference in the outcome between the categories of the dichotomous predictor increasing/decreasing as the quantitative predictor increases/decreases. This model used standardized immorality ratings as the outcome variable and was designed to predict the outcome from standardized *disgust*, *grossed-out*, and *anger* ratings at the participant level and to predict the outcome from violation type at the story level.

Contrary to the domain theory hypothesis, *grossed* ratings did not predict a significant amount of variability in immorality ($\beta = -.17$, $t(24.18) = -1.15$, $p = .262$); they were dropped from the final model. According to the resulting final model, mean immorality ratings were higher for non-health than health purity violations ($t(31.72) = 4.87$, $p < .001$); this difference increased as anger ratings increased ($t(200.09) = 2.29$, $p = .023$). Disgust, but not anger, predicted immorality ratings ($\beta_{disgust} = 0.69$, $t(19.54) = 5.22$, $p < .001$; $\beta_{anger} = 0.12$, $t(56.31) = 1.48$, $p = .142$). The correlation between immorality and *disgust* was stronger for non-health than for health purity violations ($t(11.82) = 2.56$, $p = .025$). Immorality ratings did not appear as correlated with anger in the model including both anger and disgust ratings as predictors; which can be due to both anger and disgust overlapping to a great extent in the portion of variability of the outcome they help to explain. This notion is supported by the magnitude of the regression weight obtained for anger when disgust is dropped from the model ($\beta = .47$, $t(16.57) = 8.12$, $p < .001$) and the raw correlation between these two predictors ($r_{anger-disgust} = .43$). Besides, the difference on how immoral non-health and health purity violations were rated increased with anger. On the other hand, the more disgusting was the act, the more it was rated as immoral, especially for non-health violations. In short, the degree to which the act elicited feelings of disgust and anger overlap in predicting how immoral it was, but *grossed-out* had no significant

effect on how immoral it was rated. Besides, the difference in how immoral health and non-health purity violations were rated increased with anger. To illustrate the relation of anger to immorality graphically, Figure 2 shows mean *anger* and mean immorality ratings across the 13 stories. Although inferential tests are not appropriate for these data, the relationship shown was tested from non-aggregated data by MLM. For more correlational analysis see Supplement S1-4.

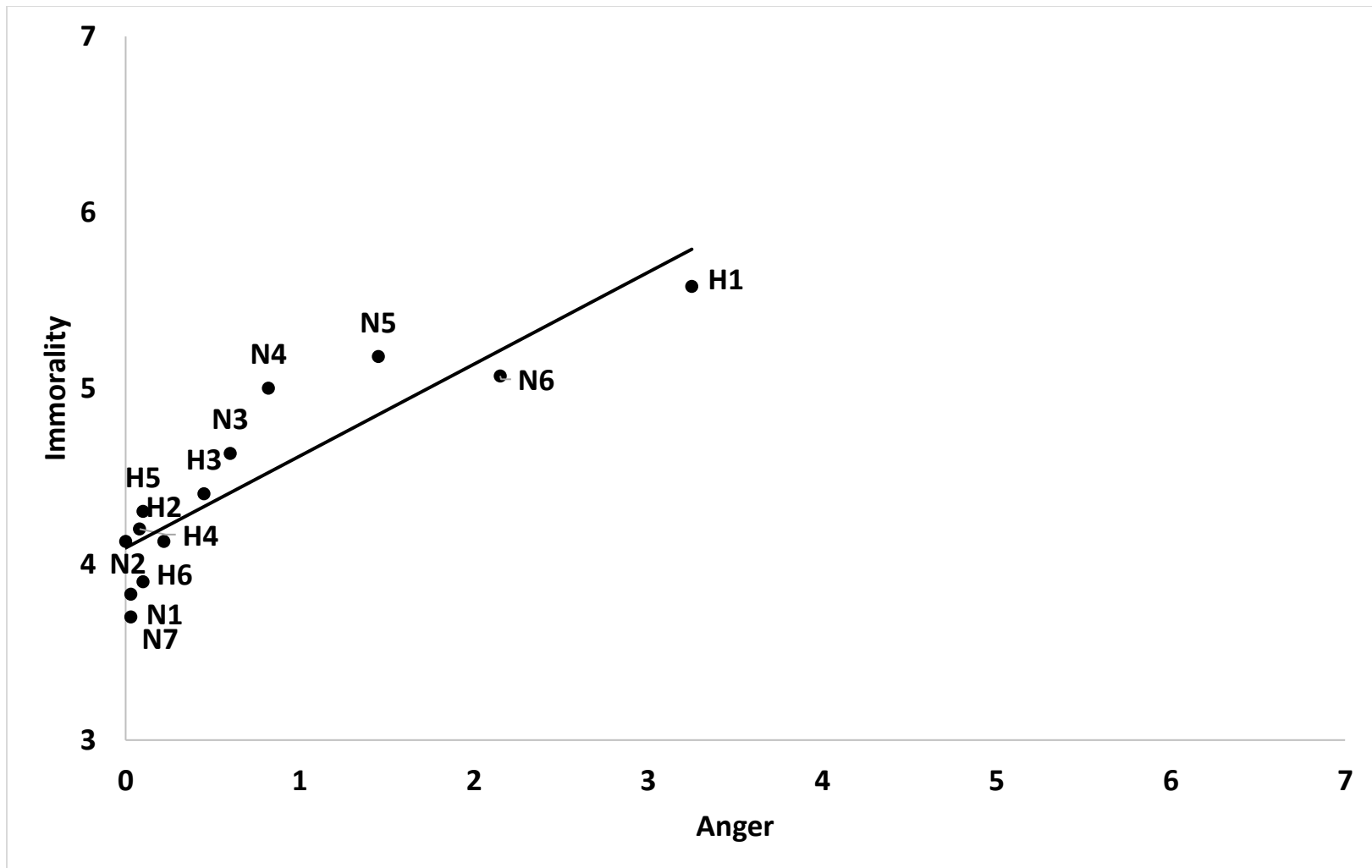


Figure 2. Immorality predicted from *anger*, Study 1. Both immorality and *anger* ratings are average ratings for the story ($N = 40$ per story). N1 = Non-Health purity violation story 1, and H1 = Health purity violation story 1.

Summary. Thirteen actions – all based on ones offered in previous studies as violations of moral purity – failed consistently to support the hypotheses made in postulating that the purity domain is homogeneous. Of the 80 participants, 45% denied any tainting of the soul to any of these alleged purity violations. When a bipolar moral-immoral rating scale replaced the usual unipolar scale typical of previous studies, some alleged examples of purity violations were judged as at least somewhat immoral, but most as either neutral or only barely immoral. Several were judged as moral.

Especially important was that health and non-health purity violations did not show the same pattern of predicted response in the witness. These two types of supposed purity violation elicited different judgments in the witness. Of the 80 participants, 51% thought non-health violations tainted the soul, whereas only 31% thought health violations did so. Thus, 69% of participants *denied* that health purity violations tainted the soul. Physical-health-related purity violations as a group elicited *grossed-out* as the modal emotion. Non-health purity violations, in contrast, elicited emotions to a minimal degree, *grossed-out* least of all, although these emotion ratings were difficult to interpret because of floor effects.

Unlike *grossed-out*, both anger and disgust were associated with immorality. However, the portion of variance of immorality they help to predict overlaps to a great extent, making it difficult to assess the relative contribution of each one of them by means of correlational methods such as partial regression weights estimation.

Study 2: Is Purity Homogeneous?

Is it Distinct from Autonomy?

Study 1 offered several findings that challenged the theorized moral domain of purity. Perhaps the most surprising finding was that most alleged purity violations, even though used in

previous studies as exemplars of a moral violation, were judged barely immoral if at all.

Although less surprising, equally important were the results indicating that purity violations are not homogeneous, but witness judgments fall into two patterns, separating health-related from non-health-related. To continue to explore these findings, we studied a new set of violations based on a set of purity norms stated by Vasquez et al. (2001), which also allowed us to include a set of autonomy violations – thus extending our study to distinctness.

Conceivably, the low immorality ratings were revealed because we replaced the typical unipolar immorality rating scale with a bipolar morality-immorality scale. In this study, we could examine not only new exemplars of purity violations but also the use of the bipolar immorality scale with autonomy violations.

The low immorality ratings in Study 1 conceivably occurred if a high proportion of its participants were liberals. Unfortunately, political orientation was not assessed. Therefore, in Study 2, we examined whether participants who identify themselves as politically conservative would judge purity violations as more immoral, disgusting, and gross than do participants who identify themselves as politically liberal -- as suggested by previous research (Graham et al., 2009).

The design was within-subjects. Each participant read 13 stories, each about an action. One was a filler story describing a highly moral action. Four stories presented autonomy violations, four health purity violations, and four non-health purity violations. Participants rated how each action made them feel on scales of *disgust*, *grossed-out*, *anger*, and two comparison emotions. Participants also rated how much the act was moral or immoral.

Method

The method is the same as in Study 1 except as noted.

Participants. Participants ($N = 96$) were American native speakers of English recruited through Amazon Mechanical Turk (46 women, 50 men; $M_{\text{age}} = 33.94$ years, range: 20-72 years). Based on the effect size in Study 1 ($d = .54$), we found that a minimum of 29 participants were needed for the conventional power value of $(1-\beta) = .80$, two tailed, .05 alpha. We recruited over 90 participants in view of the three planned contrasts (*anger vs. disgust* and *disgust vs. grossed-out* in health and non-health violations, and *anger vs. disgust* in non-health and autonomy violations). For details of the power analysis see Supplement S2-1. All participants passed an attention check by providing a unique code at the end of the survey and by typing the name of their favorite color; and so all data collected are reported here. Study 2 used a similar design to that of Study 1, in which each participant responded to sets of stories of different violation types.

Stories. The 8 purity violations (4 physical health and 4 non-health) and 4 autonomy violations were actions based on norms provided by Vasquez et al. (2001). There was also one filler story (see Supplementary Procedures 2).

Political self-identification. Participants were instructed: "Please indicate your political ideology using the scale below." Political self-identification was reported on a 6-point scale: Strongly liberal, somewhat liberal, slightly liberal, slightly conservative, somewhat conservative, and strongly conservative.

Procedure. Participants completed the survey online. The participant first read the consent form, then the general instructions, and then the thirteen stories. Participants read the filler story first. The order of the other twelve stories and the order of the emotion labels in the emotion response scale were randomized separately for each participant. After reading each story, the participant completed the emotion response scales and then the immorality scale, in

that order. After responding to all 13 stories, participants indicated their political self-identification and answered demographic questions about their sex and age.

Results and Discussion

Check on the bipolar morality-immorality rating scale. Participants' responses to our one moral filler story and autonomy violations indicated that the bipolar immorality rating scale (on which 4 indicated neutral) worked well. Participants judged the action in the moral filler story as between *moderately* and *extremely moral* ($M = 1.57, SD = 0.83$) and those in the autonomy violation stories, on average, as between *moderately* and *extremely immoral* ($M = 6.72, SD = 0.53$). (See Supplement S2-2 for ratings for each of the measures for each of the stories).

Emotional responses. The theory that purity and autonomy are qualitatively distinct domains predicts that purity – whether health or non-health – is based on disgust (especially as assessed with ratings of *gross*), whereas autonomy is based on anger. In contrast, results from Study 1 led us to anticipate a more complex pattern of emotional responses across violation types. Figure 3 presents mean ratings for each of the three most relevant emotion terms (*disgust, grossed-out, anger*) separately for health purity, non-health purity, and autonomy violations based on mean scores for the four stories within each violation type. (Results from the other emotions are presented in Supplement S2-3). Visually, the three types of violations elicited distinct emotional patterns; this appearance was confirmed by a violation type (3) x emotion (3) Analysis of Variance (ANOVA), $F(4, 380) = 139.89, p < .001$, partial eta squared .60. Autonomy violations showed the predicted dominant response, *anger*, but also similarly high levels of *disgust*. Collapsed across the health and non-health subtypes, purity violations overall elicited higher average *disgust* and *grossed-out* than *anger* ratings, as predicted in the classical theory of

purity. However, health and non-health purity violations dissociated in two revealing ways: the relative salience of *anger* and the distinctive role of *grossed-out*. As expected by domain theory, health purity violations were characterized by high levels of both *grossed-out* and *disgust*, and very little *anger*. In contrast, non-health purity violations were characterized by high levels of *disgust*, moderate levels of *anger*, with much lower levels of *grossed-out*. In short, autonomy (albeit more severe) was the mirror-image of (i.e., a reversed version of) health purity violations, but non-health purity elicited a different pattern, indeed little emotion at all, *gross* least of all.

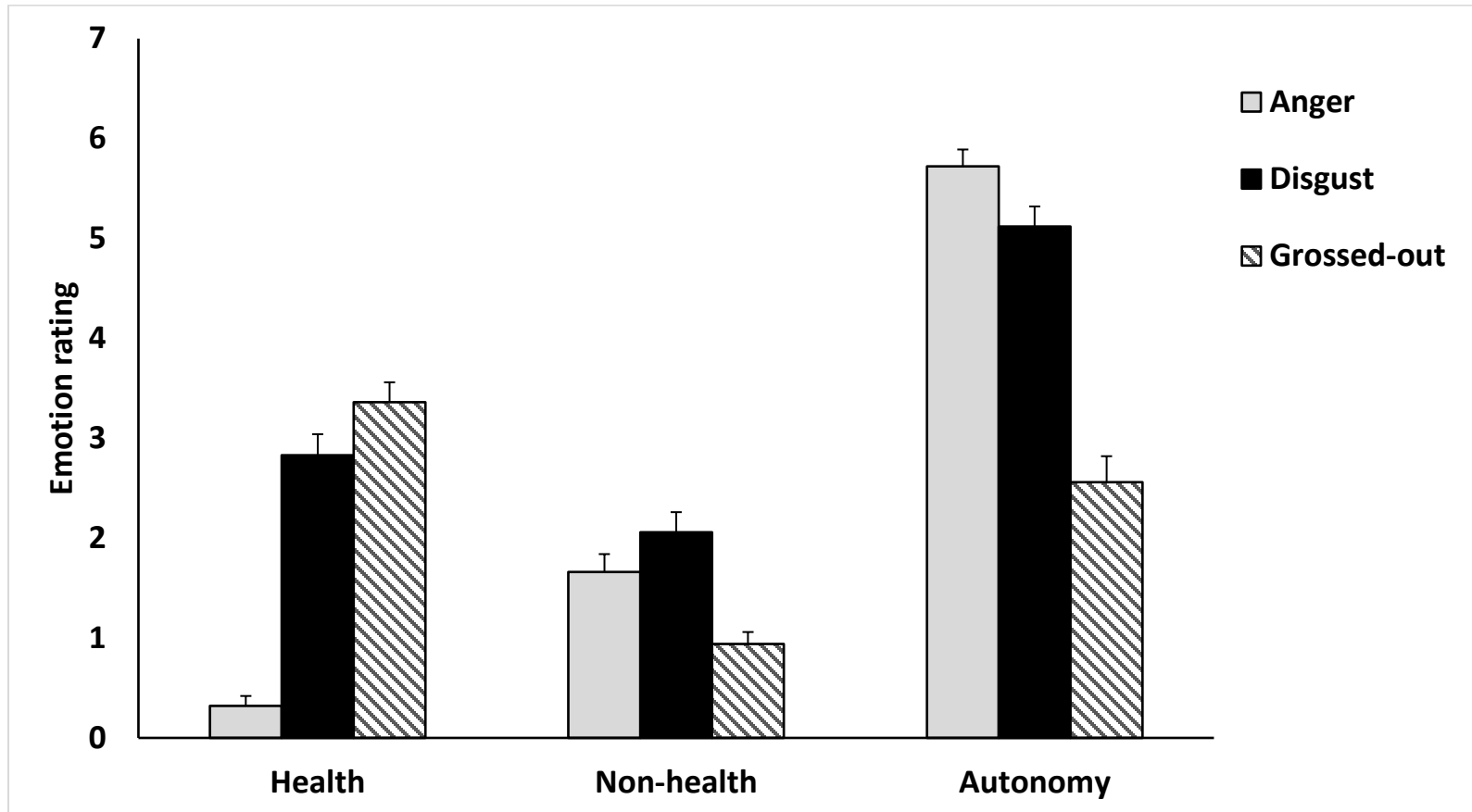


Figure 3. Mean emotion ratings for health, non-health, and autonomy violations, Study 2. Each emotion was rated on an 8-point scale with 0 defined as “no” and 7 as extreme. Standard errors are represented in the figure by the error bars attached to each column.

Further (2 x 2) ANOVAs, each with just the two purity violation types (health, non-health) and two levels of emotion (*anger* vs. *disgust* in one, and *disgust* vs. *grossed-out* in the other), confirmed the reliability of the interaction pattern shown in Figure 3 by revealing reliable interactions for *anger* versus *disgust* $F(1, 95) = 83.55, p < .001$, partial eta squared = .47, and for *disgust* versus *grossed-out*, $F(1, 95) = 61.25, p < .001$, partial eta squared = .39. This last result replicates the critical effect observed in Study 1. For non-health purity stories, due to a floor effect in the data of the present study, we also conducted a sign test, which supported the same conclusion. *Disgust* was higher than *grossed-out* for 59 out of 69 participants with 27 ties excluded), $z = 5.78, p < .001$. Finally, a follow-up ANOVA that compared non-health versus autonomy violations (i.e., excluding health violations) confirmed the predicted switch from *anger* to *disgust* as the primary emotion for non-health violations, $F(1, 95) = 37.83, p < .001$, partial eta squared = .28. Separate analyses of data from each emotion corroborate the pattern described above. These results are reported in Supplement S2-4. Evidence of generalizability was found in the means for individual stories within each violation type; means failed to identify any individual story that departed substantially from the group pattern; results are presented in Supplement S2-5.

To summarize, the emotional responses to the different violations challenged the theory that purity is homogenous and distinct from autonomy. Homogeneity of purity is challenged by the ratings of *gross*: relatively high, indeed modal, for health, negligible for non-health purity. Although autonomy violations elicited generally stronger negative emotional responses than did the other violation types, that difference is quantitative rather than qualitative. The autonomy and non-health purity violations were similar to each other in their connection to *anger* and *disgust*, but distinct from health purity violations, especially with respect to feelings of *anger*.

Morality-immorality. An assumption in the research from which we obtained our examples of purity and autonomy norms is that violations are immoral. Autonomy violations on average were judged the most immoral -- between moderately and extremely immoral. On the bipolar rating scale, in which neutral was 4.0 and maximum 7, the mean for autonomy was $M = 6.71$, $SD = 0.53$; means for individual stories ranged from 6.58 to 6.83. Non-health purity violations were intermediate—close to barely immoral ($M = 5.08$, $SD = 0.84$; means for individual stories ranged from 4.54 to 5.66). And health violations the least immoral -- indeed, on average, they were neutral ($M = 4.04$, $SD = 0.65$; means for individual stories ranged from 3.93 to 4.25). ANOVA revealed that, collapsed across the 4 stories within each type of violation, immorality ratings differed across violation type, $F(2, 190) = 462.33$, $p < .001$, partial eta squared = .83. Also, each violation type was different from each other type, $ps < .001$. The health story immorality ratings were not reliably different from the neutral 4.0, $t(95) = 0.66$, $p = .508$.

Political orientation. The prediction from previous research is that violations in the purity moral domain would be judged *disgusting*, *gross*, and *immoral* by everyone but more so by self-identified conservatives than by self-identified liberals. Presumably, if purity is a single homogenous moral domain, then this prediction holds for purity violations whether health-related or non-health related. The present results provided inconsistent support for the role of political outlook, but what support there was varied between health and non-health purity violations.

A series of multilevel mixed-effect regression models were used to predict standardized immorality and emotion ratings as the outcome variables. They were designed to predict variability between participants in the outcome variables from standardized responses in political

orientation and to predict differences between stories from the type of violation (purity non-health vs. purity health vs. autonomy).

After controlling for the type of violation, immorality ratings were correlated with conservatism ($\beta = .14$, $t(3.84) = 2.92$, $p = .046$). This correlation were higher for autonomy and non-health purity violations combined than for health purity violations ($t(3.56) = 3.30$, $p = .036$), but the difference between autonomy and both purity violations were not significant ($t(3.56) = 2.50$, $p = .074$).

Anger responses, on average, were greater for more conservative participants ($\beta = .17$, $t(9.00) = 3.41$, $p = .008$). The magnitude of the correlation between anger and conservatism was higher for purity violations than for autonomy violations ($t(9.46) = 3.94$, $p = .003$), and also higher for non-health purity and autonomy violations combined than for health violations ($t(9.46) = 2.53$, $p = .031$).

Disgust responses, on average, were also greater for more conservative participants ($\beta = .19$, $t(47.80) = 4.17$, $p < .001$). The magnitude of the correlation between disgust ratings and conservatism was stronger for both purity health and non-health combined than for autonomy violations ($t(44.85) = 2.96$, $p = .005$), and also greater for autonomy and non-health violations combined than for health violations ($t(44.85) = 4.57$, $p < .001$).

However, across all violation types, *grossed-out* responses were not reliably correlated with political orientation ($\beta = .08$, $t(75.15) = 1.61$, $p = .111$). Yet, the magnitude of this correlation was higher for both purity than for autonomy violations ($t(77.66) = 2.34$, $p = .022$) and also higher for autonomy and non-health purity violations combined than for health purity ($t(77.66) = 2.38$, $p = .020$).

In sum, immorality, anger, and disgust ratings increased as the political ideology was more conservative, especially for purity violations and, among them, non-health violations. Crossed-out was not reliably associated with conservatism.

An alternative approach in which political orientation was a dichotomous variable (conservatives, $n = 40$ vs. liberals, $n = 56$) failed to show that the conservatives, compared to liberals, found purity violations more *disgusting*, *gross*, and immoral. The mean immorality rating was not significantly greater for conservatives than for liberals in each of the violation types. Contrary to the prediction, for health violation, immorality rating was significantly greater for liberals ($M = 4.16$, $SD = 0.33$) than for conservatives ($M = 3.87$, $SD = 0.91$), $t(94) = 2.19$, $p = .031$. Immorality rating for non-health and autonomy violations did not significantly differ between liberals and conservatives: non-health, liberals ($M = 5.05$, $SD = 0.77$), conservatives ($M = 5.14$, $SD = 0.93$), $t(94) = -0.51$, $p = .613$; autonomy, liberals ($M = 6.74$, $SD = 0.51$), conservatives ($M = 6.68$, $SD = 0.56$), $t(94) = 0.58$, $p = .560$.

In a 3 (emotions) x 3 (violation type) x 2 (political orientation) ANOVA, the three way interaction was not significant, $F(4, 376) = 1.79$, $p = .129$, partial eta squared .02. The non-significant result is clarified by the pattern of emotion ratings across violation types in Figure 4 which show how similar the results for conservatives match those from liberals.

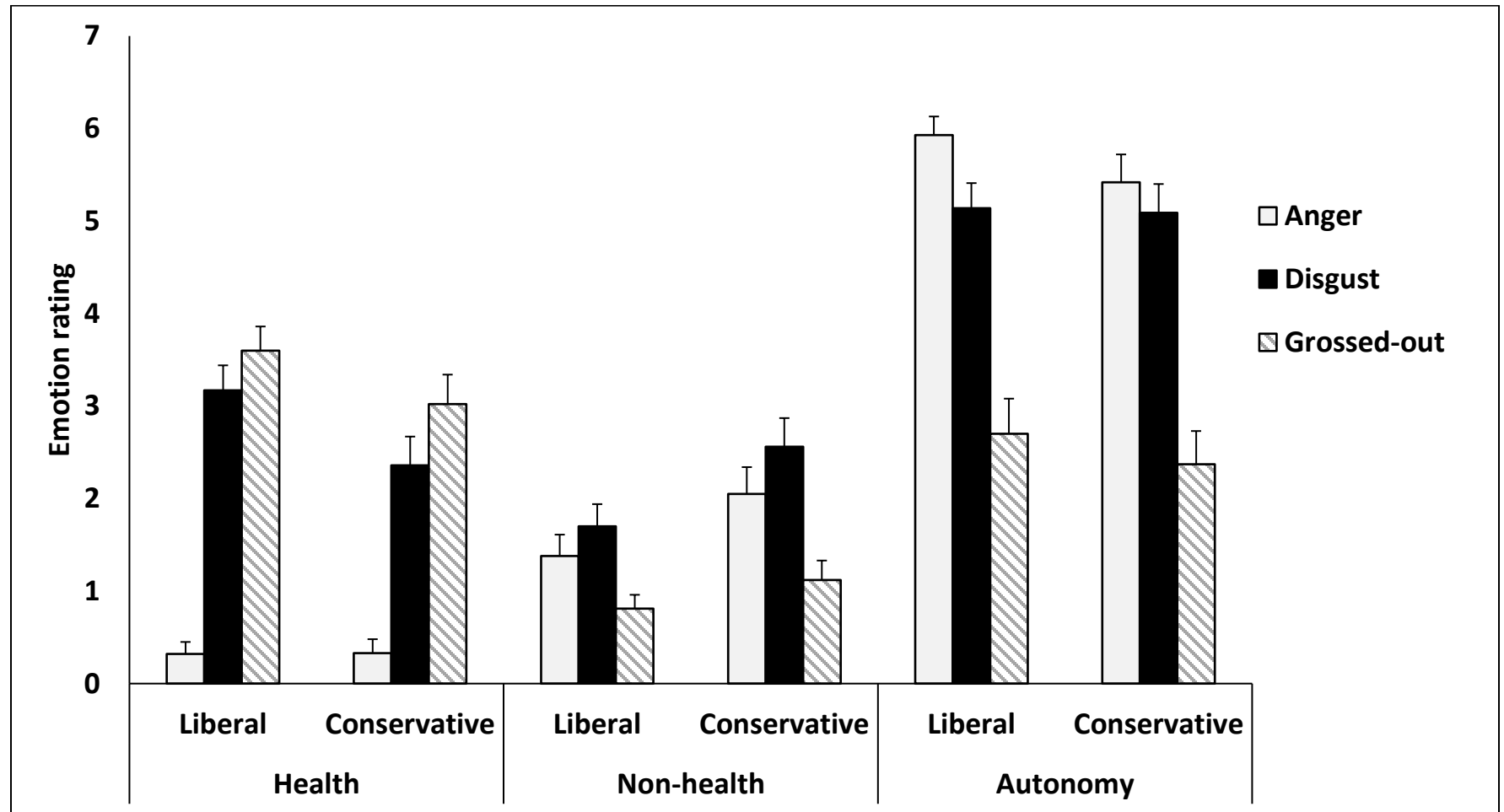


Figure 4. Mean emotion ratings for health, non-health, and autonomy violations for liberal ($n = 56$) and conservative ($n = 40$) participants, Study 2. Standard errors are represented in the figure by the error bars attached to each column.

The question is whether a theoretical claim for a homogeneous purity domain can be rescued by restricting it to conservatives. This rescue remains a hypothetical possibility perhaps but, at least according to the present findings, a weak one of a correlational nature and only for non-health violations. Compared to more liberal participants, more conservative participants found non-health purity violations more *disgusting* and *angering*. A comparison on average mean ratings failed to show that the conservatives, compared to liberals, found purity violations more *disgusting*, *gross*, or immoral.

Continuum across violation-types. Even though our previous analyses highlighted distinct pattern of participant reactions across the three violation types, other findings also indicate that at least some part of the distinct reactions may result from the differences in immorality across the violation types. Health violations were the least immoral, non-health violations were intermediate, and autonomy violations the most. Anger ratings follow the same pattern: Anger ratings were the least for health violations, intermediate for non-health violations, and the most for autonomy violations. The immorality ratings and the anger ratings are consistent with a theory in which purity and autonomy violations fall on a continuum rather than into qualitatively distinct clusters.

This theory was explored with a multilevel mixed-effect regression model with standardized immorality ratings as the outcome variable. The model was designed to predict variability in outcome among participants from the standardized emotion ratings and differences in outcome between stories from the violation type. Contrary to domain theory, *grossed-out* ratings were not correlated with immorality ratings ($\beta = .03$, $t(11.40) = 0.61$, $p = .557$) and were dropped from the final model.

According to the final model, there was a positive correlation between immorality and both anger and disgust ratings ($\beta_{anger} = .40$, $t(2.65) = 4.64$, $p = .025$; $\beta_{disgust} = .40$, $t(3.79) = 7.46$, $p = .002$). The correlation between immorality and disgust ratings was also stronger for purity violations (combined health and non-health) than for autonomy violations ($t(3.51) = 4.30$, $p = .017$), but the difference between non-health purity and autonomy violations combined and health purity did not reach statistical significance ($t(2.96) = 3.19$, $p = .051$). No significant differences between types of violation were found in the magnitude of the correlation between immorality and anger (autonomy vs. both purity: $t(2.70) = 1.01$, $p = .393$; health purity vs. non-health purity and autonomy combined: $t(3.99) = 0.73$, $p = .505$). No significant effect of the type of violation was found in mean immorality ratings either (autonomy vs. both purity combined: $t(2.54) = 3.42$, $p = .054$; purity health vs. autonomy and non-health purity: $t(2.69) = 2.63$, $p = .088$).

In other words, the more immoral a violation was rated, the more angering and more disgusting it was rated. The correlation between immorality and disgust was stronger in the case of health and non-health purity violations than in autonomy violations. In contrast, the correlation between immorality and anger did not differ across violation types. There were no significant differences in how immoral the types of violations were rated on average. To illustrate the relation of anger to immorality graphically, Figure 5 shows mean *anger* and mean immorality ratings across the 12 stories. Although inferential tests are not appropriate for these data, the relationship shown was tested from non-aggregated data by MLM.

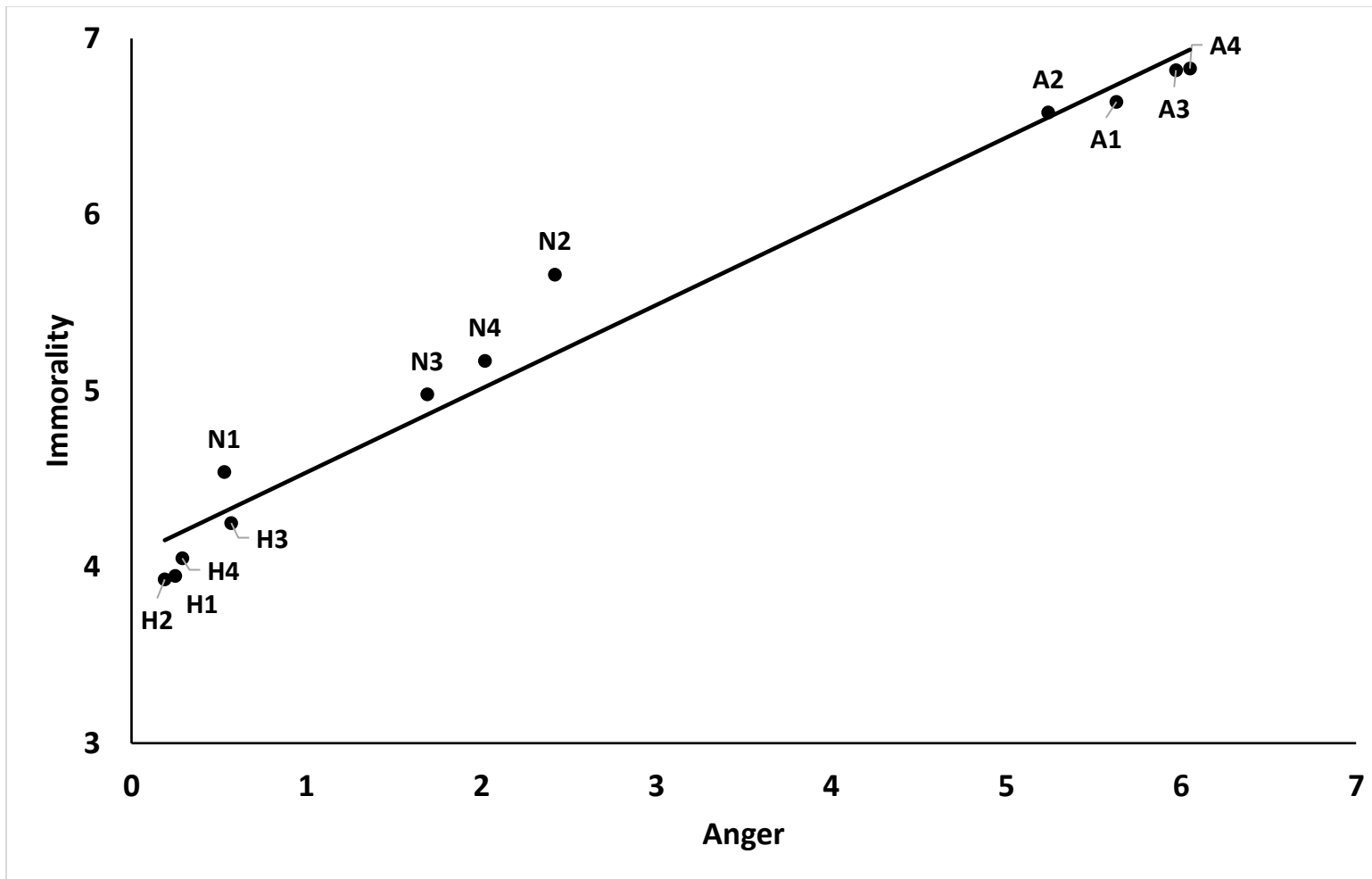


Figure 5. Immorality predicted from anger, Study 2. Immorality was measured on a bipolar rating scale, in which neutral was 4.0.

Both immorality and anger ratings are average ratings for the story ($N = 96$ per story). A1 = Autonomy violation story 1, N1 = Non-Health purity violation story 1, and H1 = Health purity violation story 1.

This finding offers a new perspective on the apparent separation of domains by emotion such that autonomy violations elicit *anger*, whereas purity violations do not. For both domains, *anger* predicts immorality. The appearance of *anger* being irrelevant in purity arises because both *anger* and immorality are low in that domain; nevertheless, within purity, the more *angering* the violation, the more it was judged immoral.

Rottman et al. (2014) similarly reported that the pattern of mean emotion responses was not exactly that predicted by the classic theory of purity versus autonomy. Rottman et al. suggested an alternative: that correlations based on individual differences within each story provide better indices of how emotion relates to judged immorality. We therefore examined such correlations in the present data. These and other results are reported in Supplement S2-6.

Summary. On our bipolar scale of immorality, violations of purity norms based on prior research elicited mild judgments (on average, neutral or between neutral and “barely”). This finding is important because our bipolar morality-immorality scale credibly captured participants’ judgments of immorality: On that scale, on average, actions of a person aiding another were judged as moral, whereas actions of a person harming another person were judged as immoral. Thus, the alleged immorality of purity violations examined in the literature is a question greatly in need of study.

Immorality judgments of violations *within* the purity domain were also importantly different. On average, non-health purity violations were judged as on the immoral side of the scale, whereas the health-related purity violations were judged as neutral – neither moral nor immoral -- on that same scale. Differences within purity were also found with respect to feelings of *disgust* and *grossed-out*: *disgust* and *anger* were used for non-health violations, *grossed-out* for health violations.

Political orientation was linked to judgments of non-health violations (such as showing disrespect for religious objects) in that conservatives reacted more negatively, but the predicted dominance of judgments of feeling *grossed-out* and finding them immoral was not well supported.

Mean ratings for the 12 stories revealed an interesting pattern: whatever the domain, *anger* ratings predicted a relevant amount of the variation in immorality, a finding that suggests a common factor across putative violation domains. This finding also provides an important perspective on reported qualitative differences in the average emotion seen in the different domains: low *anger* ratings in health-related purity violations are consistent with low immorality ratings.

Study 3:

Replication and Extension: Tainting the Soul

Study 2 raised serious questions about the theory of purity as a homogeneous moral domain qualitatively distinct from the autonomy domain. Even though the exemplars used in Studies 1 and 2 as violations had been based on exemplars and explicit moral norms stated in previous studies, we had concerns. First, we had relied on a modest number of violations, and so in this Study 3 we used more stories, none of which were included in Studies 1 or 2. The larger number of stories allowed a more robust examination of our findings on how immorality is predictable from emotion across domains.

Second, the purity violations studied so far were judged neutral or barely immoral, and we sought to include more severe ones, but that, nonetheless, had been used in previous studies as purity violations. Because it is questionable whether the domain of moral purity can be properly examined with actions that are not perceived as immoral, we used our own judgment to

exclude any purity violation that seemed on the moral side of the ledger or close to it. In this way, we sought a set of more severe purity violations. One obstacle was that many purity violations rated as highly immoral were related to sex, and so for a non-health violation, we included stories that portrayed sex but in actions -- watching pornography or fantasizing-- that are not literal threats to physical health.

And third, Study 3 used a new measure of tainting the soul – purported to indicate violations specific to purity. Our Study 1 had borrowed a scale from Rottman et al. (2014) for tainting the soul. However, our results showed a highly skewed response distribution and floor effect. This result raised the problems mentioned earlier with unipolar response formats, and our Study 2 did not include a scale for tainting the soul. Study 3 introduced a bipolar format for tainting the soul.

To ensure robustness of results across differences in method, there were other changes in details. In contrast to Studies 1 and 2, which had an entirely within-subjects design, Study 3 made both violation type and story between-subjects factors. We made a slight change in the rating scale for emotion items. We also included a manipulation check for the health purity, non-health purity, and autonomy violations. Prior to receiving the experimental story, the participant was given two filler stories, one concerning an action generally judged moral (donating to charity) and another something immoral (embezzling) but in the alleged community rather than purity or autonomy domains.

Method

The method is the same as in Study 1 except as noted.

Participants. Participants ($N = 1312$) were Americans recruited through Amazon Mechanical Turk (747 women, 565 men; $M_{\text{age}} = 35.68$ years, range: 18-87 years). One reason for

recruiting a large sample was to introduce additional screening to ensure that participants were responding seriously in their judgments of morality. An additional 478 participants were recruited and participated but their data were excluded from the data analyses reported here because they failed our test of paying appropriate attention and seriousness as seen in their responses to the filler stories (336 participants judged the moral filler story as immoral, another 116 participants judged the immoral filler story as moral, and 26 made both types of anomalous responses). In Supplement S3-1 is presented an analysis in which their data were included ($N = 1790$). The patterns reported there were all similar but weaker. We based our power analysis on Study 2 (for details of the power analysis see Supplement S3-2). For a medium to large sized effect ($d = .62$), with the conventional power value of $(1 - \beta) .80$, two tailed, $.05$ alpha, the adequate sample size was 42 participants per story. This is a conservative approach because our primary hypotheses in Study 3 were tested collapsing across 9 different stories for each violation type, with a sample size of approximately $n = 50$ (range: 36 – 66) for each story. To achieve this size, we aimed to recruit at least 15 participants for each version of each story. (Each story had four versions with a male or female protagonist and with a random or reverse order of dependent variables).

Stories. There were 27 stories describing actions based on events offered in the literature. There were 18 purity violations (9 physical health and 9 non-health) and 9 autonomy violations. For each of the stories, sex of protagonist (male vs. female) was a between-subject factor. All 27 experimental stories and the justification for choosing each story are given in Supplementary Procedures 3.

Unipolar emotion response scales. The emotion response scales included 7 emotions, all in an adjectival form: *angry*, *grossed-out*, *disgusted*, *contemptuous*, *sad*, *happy*, and *proud*.

The score for each emotion was 0 for “no” and the intensity rating for “yes.” The range was 0 to 6.

Bipolar scale for tainting the soul. Participants were asked, for example, for the female version: “Did the person’s actions affect her own soul (her spiritual side, her relation to what is sacred?).” Participants could choose by clicking 1-“extremely tainting to her soul”, 2-“moderately tainting”, 3-“barely tainting”, 4-“neither tainting nor beneficial”, 5-“barely beneficial”, 6-“moderately beneficial”, 7-“extremely beneficial to her soul.” Thus, the range was 1 to 7.

Design. The design was mixed. There were 9 levels of story nested within the three violation types; the three were between-subjects factors, along with 2 levels of sex of protagonist and 2 levels of order of dependent variables. Each participant responded to one of the 108 versions of the questionnaire. Each participant rated each story on 7 emotions as well as immorality and tainting of the soul. Our analyses did not include sex of protagonist or order of dependent variables.

Manipulation check. The design of the manipulation check mirrored that of the study itself by including the same general instructions and filler stories. There were 540 participants who did not participate in the main study and who were recruited through Amazon’s Mechanical Turk and compensated for their work (347 women, 253 men; $M_{\text{age}} = 34.74$ years, range: 18-85 years). There were two dependent variables: threatens own health and harms another. Health purity violations were clearly judged as actions that threaten one’s own health rather than harm another person. In contrast, autonomy violations were judged as actions that harm another person rather than threaten one’s own health. Non-health purity violations were judged neither as threatening one’s own health nor as harming another person (see Supplement S3-3 for details).

Results and Discussion

Check on the bipolar morality-immorality rating scale. Participants' responses to the filler stories showed that our bipolar morality-immorality rating scale worked well. (See Supplement S3-4 for ratings for each of the measures for each of the stories). On a scale in which 4 is the tipping point between moral and immoral, the subset of 1312 participants who passed screening responded as expected to the immorality of the positive filler item ($M = 1.71$) and negative filler item ($M = 6.56$). The distinction did not vary significantly as a function of the violation type of other stories seen by the participants.

Emotional responses. As in Study 2, we tested the domain theory hypothesis that purity violations -- whether health related or not -- elicit disgust, whereas autonomy violations elicit anger. Figure 6 presents mean ratings for each of the three emotions of most interest (*disgust*, *grossed-out*, *anger*) separately for health purity, non-health purity, and autonomy violations based on average scores for 9 stories within each violation type. (Assessment of consistency across the 9 stories within each violation type is reported in Supplement S3-5. Results from the other emotions are presented in Supplement S3-6). The pattern closely matched that described above for Study 2. The three types of violations did not elicit the same emotional response, as confirmed by a type x emotion interaction, $F(4, 2618) = 285.92, p < .001$, partial eta squared = .30. Autonomy violations showed the predicted dominant response: *anger*, but also similarly high levels of *disgust*. Collapsed across the health and non-health subtypes, purity violations overall elicited higher average *disgust* and *grossed-out* than *anger* ratings, as predicted by domain theory. However, as in Study 2, health and non-health purity violations dissociated based on the relative salience of *anger* and *grossed-out*. Consistent with domain theory, health violations were characterized by high levels of *grossed-out*, less *disgust*, and very little *anger*. In

contrast, non-health violations were characterized by relatively high levels of *disgust* with lower levels of *grossed-out* and *anger*. In short, autonomy (albeit again more severe) was roughly the mirror-image of (reverse of) health purity violations, but non-health purity elicited a different pattern, indeed less emotion overall, especially *gross*. Autonomy violations elicited greater feelings of disgust and grossed-out than did non-health purity violations.

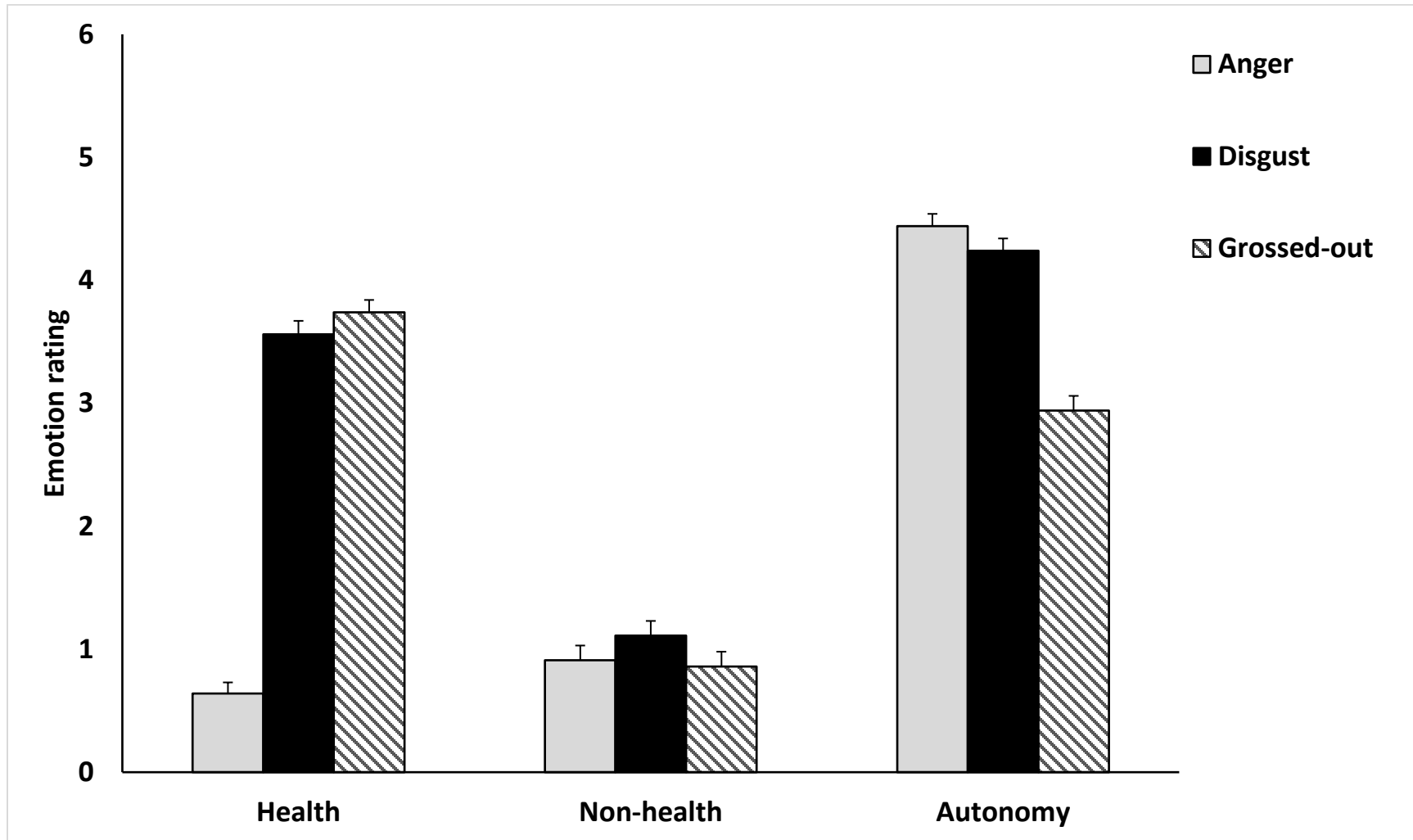


Figure 6. Mean emotion ratings for health, non-health, and autonomy violations, Study 3. Each emotion was rated on an 7-point scale with 0 defined as “no” and 6 as extreme. Standard errors are represented in the figure by the error bars attached to each column.

Follow-up ANOVAs confirmed the reliability of the interaction pattern in Figure 6. An ANOVA with two purity violation types (health, non-health) and two levels of emotion (*anger* vs. *disgust*) revealed a reliable interaction, $F(1, 858) = 210.34, p < .001$, partial eta squared = .20. Similarly, an ANOVA with two purity violation types (health, non-health) and two of emotion (*disgust* vs. *grossed-out*) also yielded a reliable interaction, $F(1, 858) = 36.26, p < .001$, partial eta squared = .04. For non-health purity stories, due to a floor effect in the data of the present study, we also conducted a sign test, which supported the same conclusion. *Disgust* was higher than *grossed-out* for 82 out of 126 participants with 295 ties excluded, $z = 3.30, p = .001$. Finally, a follow-up ANOVA that compared non-health versus autonomy violations (i.e., excluding health violations) confirmed the predicted switch from *anger* to *disgust* as the primary emotion for non-health violations, $F(1, 871) = 41.26, p < .001$, partial eta squared = .04. These details replicate the relevant effects observed in Studies 1 and 2.

To assess the type x emotion interaction from a different perspective, we analyzed ratings for *anger*, *disgust*, and *grossed-out* separately to evaluate predicted distinctions between purity and autonomy and between health and non-health purity violations. These analyses are reported in Supplement S3-7 and were consistent with the patterns described above.

In sum, witnesses' emotional responses to violations did not point to homogeneity of purity violations or to a sharp distinction between purity and autonomy. Participants dissociated health purity violations from non-health purity violations: Health purity violations elicited mainly *grossed-out* and *disgust* and little *anger*. In contrast, non-health purity violations elicited broadly comparable levels of *disgust*, *grossed-out*, and *anger*. In eliciting *anger* and *disgust*, non-health purity violations more resembled less severe autonomy violations than health purity

violations. Importantly, on average, autonomy violations elicited more feeling of being grossed-out than did non-health purity violations, even though grossed-out is likely closest to the hypothesized emotional basis of purity norms.

Tainting of the soul. Tainting the soul was offered by Rottman et al. (2014) as a marker of a purity violation. With our bipolar tainting scale (after reverse coding, with 1 as benefitting the soul, 4 as neutral, and 7 as the greatest amount of tainting), autonomy violations were judged here as tainting the soul the most—between moderately and extremely tainting ($M = 6.21$, $SD = 1.32$); non-health purity violations were intermediate— between barely and moderately tainting ($M = 5.33$, $SD = 1.50$); and health purity violations were the least tainting-- between neutral and barely tainting ($M = 4.44$, $SD = 0.99$). The main effect of violation type was strong, $F(2, 1309) = 210.57$, $p < .001$, partial eta squared = .24. Bonferroni comparisons confirmed that each violation type was significantly different from each of the two other types. Proportion of participants who found the violations as tainting the soul also showed the same pattern: 124 of 439 participants found health violations tainting the soul to some degree (<4); 260 of 421 participants found non-health violations tainting the soul to some degree; and 380 of 452 participants found the autonomy violations as tainting the soul to some degree. The ordering of violation types is the opposite of that implied by thinking of tainting the soul as a marker of a purity violation; however, note that the order might be altered by varying the severity of infraction across the violation types.

Further analyses question the use of taint as a marker of purity violations alone. A multilevel mixed-effect regression model was fit using standardized tainting to the soul ratings as the outcome variable. It was designed to predict variability between participants in outcome from standardized immorality ratings while controlling for the effect of the type of violation (purity

non-health vs. purity health vs. autonomy) in outcome between stories. According to this model, there were a positive correlation between tainting to soul and immorality ratings ($\beta = .71$, $t(48.90) = 18.80$, $p < .001$), which was greater in autonomy than in both purity violations combined ($t(1258.95) = 2.30$, $p = .022$) and also in purity health violations compared with autonomy and non-health violations combined ($t(1340.77) = 2.63$, $p = .009$).

Morality-immorality. Are purity violations immoral? Despite our efforts to include only clearly immoral purity violations, health purity violations were judged as barely immoral. That finding replicated Studies 1 and 2. On average, the autonomy violations were judged the most immoral -- between moderately and extremely ($M = 6.43$, $SD = 1.18$; means for individual stories ranged from 5.92 to 6.79); non-health purity violations were intermediate—between barely and moderately immoral ($M = 5.38$, $SD = 1.39$; means for individual stories ranged from 4.29 to 6.58); and health purity violations the least immoral--between neutral and barely immoral ($M = 4.50$, $SD = 1.08$; means the individual stories ranged from 4.24 to 4.75). Mean ratings on the morality-immorality scale for each of the stories are provided in Supplement S3-4. For health purity violations, none of the nine stories were more than barely immoral. The main effect of violation type was strong, $F(2, 1309) = 279.51$, $p < .001$, partial eta squared = .30. Bonferroni comparisons confirmed that each of the three violation types was significantly different from each of the other two.

Continuum across violation-types. In Study 2, analyses offered an interesting perspective on the relation of emotion to immorality—purity and autonomy are not distinct domains but may lay on a continuum. Still, those analyses require replication. Fortunately, the pattern of results here replicated the one from Study 2 with a larger sample of (different) stories.

A multilevel mixed-effects regression model was fit using immorality ratings as the outcome variable. The model was aimed at predicting the variability in outcome between participants from anger, disgust, and grossed-out ratings while controlling for the effect of the type of violation in outcome between stories. As in the previous studies, and despite the differences in grossed-out ratings between violation types (see Emotional responses section), grossed-out ratings were not correlated with immorality ratings and, therefore, they were dropped from the final model ($\beta = .10, t(37.70) = 1.61, p = .116$). According to the final model, immorality ratings were correlated with both anger and disgust ($\beta_{anger} = .16, t(866) = 3.06, p = .002$; $\beta_{disgust} = .49, t(25.69) = 7.66, p < .001$). The correlation between immorality and anger was stronger for stories describing autonomy than both purity violations ($t(1294.24) = 2.10, p = .036$), whereas the opposite happened for the correlation with disgust ($t(1285.43) = 3.36, p = .001$). The correlation between immorality and anger was higher for health purity than for both autonomy and non-health purity violations combined ($t(1241.60) = 4.32, p < .001$), whereas the opposite was found for the correlation with disgust ($t(1292.03) = 5.28, p < .001$).

In other words, the more immoral a violation was rated, the more angering and more disgusting it was rated. To illustrate the relation of anger to immorality graphically, Figure 7 shows mean *anger* and mean immorality ratings across the 27 stories. Although inferential tests are not appropriate for these data, the relationship shown was tested from non-aggregated data by MLM.

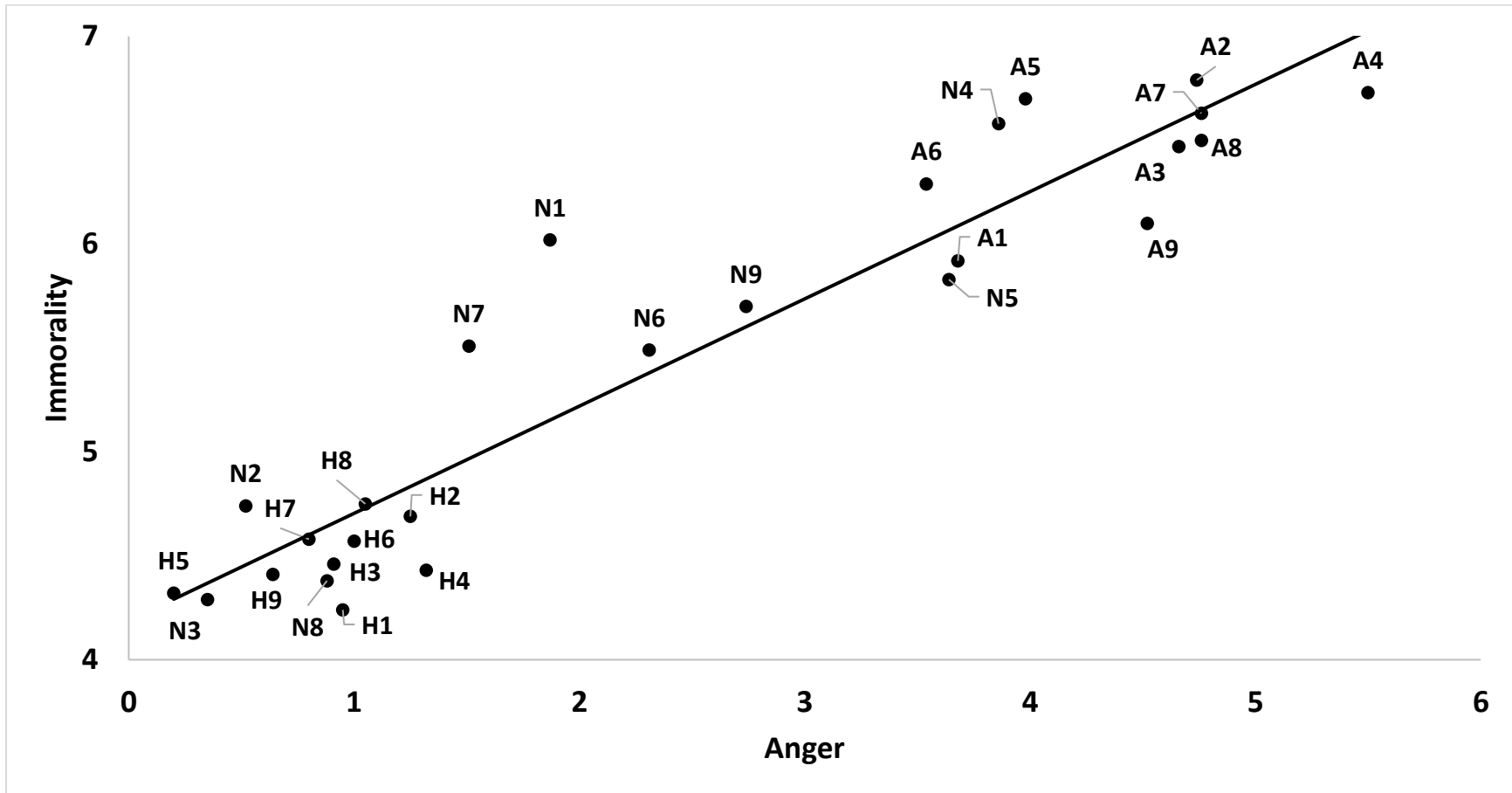


Figure 7. Immorality predicted from anger, Study 3. Immorality was assessed on a bipolar rating scale, in which neutral was 4.0. Both immorality and anger ratings are average ratings for the story (*n* per story ranged from 36 to 66). A1 = Autonomy violation story 1, N1 = Non-health purity violation story 1, and H1 = Health purity violation story 1.

Testing the homogeneity of health and non-health items by cluster analysis. The domain view is that purity and autonomy are distinct domains, whereas health and non-health are the same domain. We next used a different analytic method to provide another opportunity for this traditional view to emerge. In a hierarchical cluster analysis (Between-groups linkage method based on squared Euclidean distance, SPSS Version 27, see Figure 8), we used mean *anger*, *disgust*, and *grossed-out* ratings for each story (3 data points for each story; thus 27 data points (3 x 9) within each violation type). These ratings are traditionally hypothesized to distinguish different domains. To plot the results, ratings of *grossed-out* and *anger* (rather than *disgust*) were chosen as two axes to maximize the distinction between purity and autonomy domains. (The corresponding dendrogram is presented in Supplement S3-8.) First, the stories did not fall in the space as domain theory would anticipate: that is, purity violations would have been expected to be uniformly low on anger but range along the grossed-out axis (i.e., form a vertical column). The autonomy violations would have been expected to be uniformly low on grossed-out but range along the anger axis (i.e., form a horizontal line). Second, the tightest clusters of most similar items are indicated in solid ellipses. The more inclusive clusters indicated in dashed ellipses are relevant to the putative three domains. There was a distinct set of autonomy items which are high on the anger dimension, a tight set of health items, which are high on the grossed-out dimension, and a mixed set of health and non-health items which are low on both the anger and grossed-out dimensions. There was a more inclusive cluster of both non-health and autonomy items, high on both anger and grossed-out. There was a tight cluster of non-health and autonomy items, high on anger, moderately low on grossed-out. Consistent with other results, these groupings of items offer little clear support for the distinctiveness of a purity domain.

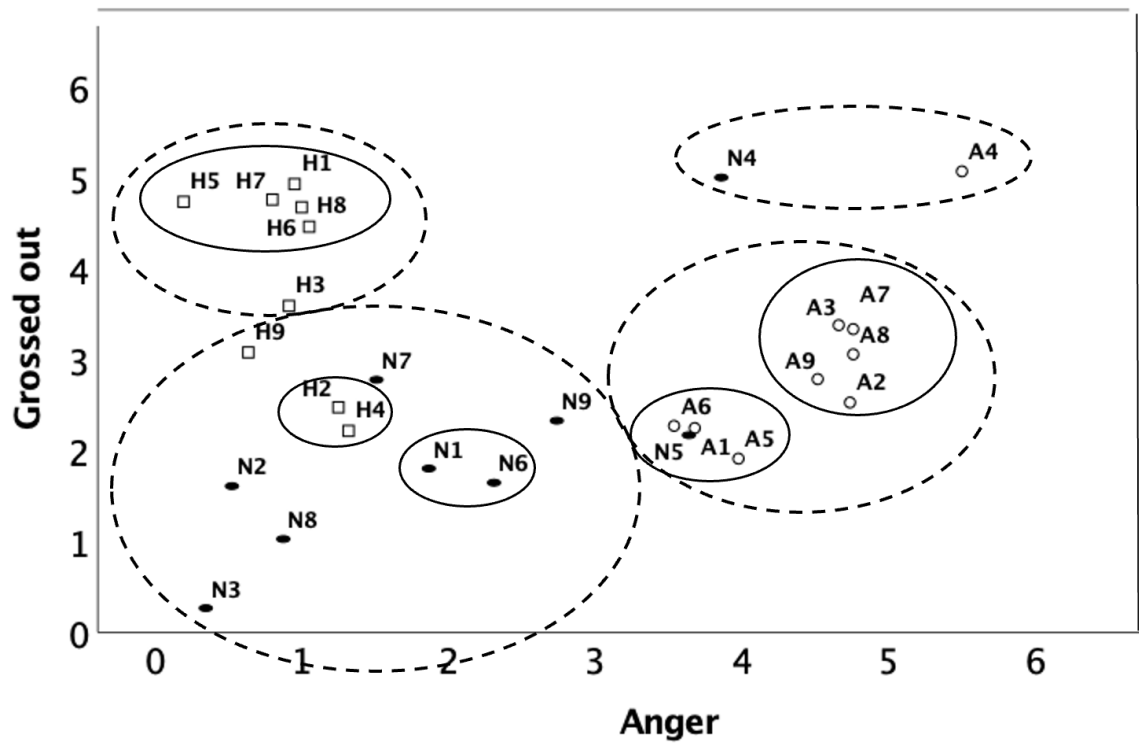


Figure 8. Twenty-seven violations plotted in a space defined by the emotion they elicited. Circles give the results of a hierarchical clustering of items based on *anger*, *disgust*, and *grieved-out* ratings. A1 = Autonomy violation story 1, N1 = Non-health purity violation story 1, and H1 = Health purity violation story 1. Solid ellipses indicate empirical groups of highly similar items, and dashed ellipses indicate empirical groups of moderately similar stories.

We report more individual difference analyses in Supplement S3-8. Across participants ($N = 1312$), immorality was positively and significantly correlated with *anger*, *disgust*, *grieved-out*, *sad*, *contempt*, and taint to soul. We again followed Rottman et al.'s (2014) suggestion to use correlations to examine individual differences within each violation type.

Summary. The basic idea of a domain theory of moral judgment is that each domain consists of a homogeneous set of violations qualitatively distinct from those in any other domain. At its core, each domain was theorized to have a discrete emotional basis, specifically, disgust for purity and anger for autonomy. The implication is that purity violations would be homogeneously disgusting, something that violations in any other domain are not. Figures 6 and 7 show that this idea fared poorly.

Furthermore, emotional profiles within the set of purity violations differed for health and non-health. For health purity violations, the primary reaction was feeling *grossed-out*, whereas for non-health purity violations, the primary response was *disgust*. Thus, *grossed-out* and *disgust* are not interchangeable. Health purity violations elicited similar *grossed-out* and *disgust*, and less anger; whereas non-health purity violations elicited similar *anger* and *disgust* – much as did autonomy violations. The autonomy domain elicited the predicted *anger*, but also *disgust* and, to lesser extent, *grossed-out*. Indeed, scores on all of these three emotions were higher in autonomy than in non-health purity. Troubling for the theory of discrete domains, *disgust* was higher for autonomy than for health purity violations. Although such main effects are suggestive, the differences in immorality between purity and autonomy renders them no more than that. Still, interaction patterns strongly supported the suggestion.

Violating a purity norm was traditionally hypothesized to taint the soul in a way that defines a purity violation. That hypothesis too fared poorly. Violating an autonomy norm was judged more tainting than violating a purity norm. And within the purity domain, violating a non-health purity norm was more tainting than violating a health purity norm – which was judged barely tainting if at all. Tainting is a marker, not of purity, but of immorality more generally.

Study 3 broadly confirmed the results of Studies 1 and 2: The mean emotion ratings for *anger*, *disgust*, and *grossed-out* indicated that the purity domain can usefully be subdivided into health and non-health violations. Health scenarios are characterized by low levels of *anger* and relatively high levels of *grossed-out*; non-health violations are characterized by more similar levels of all three emotions, with *disgust* being slightly higher than *grossed-out* and *anger*. Autonomy violations are characterized by low levels of *grossed-out*. Finally, anger ratings by themselves provide a strong predictor of immorality across allegedly distinct domains. Let us suggest that witnesses are angered by immoral deeds. Indeed, the data so far suggest that acts vary along a continuum, irrespective of theoretical domain – a continuum marked by ratings of immorality, anger, and tainting of the soul.

Study 4:

Pre-registered Replication and Extension: Controlling for Confounding Variables

Study 3 replicated our findings in Study 2—purity is not a homogeneous moral domain, nor is it qualitatively distinct from the autonomy domain. In Study 4, we aimed to replicate in a pre-registered study our findings so far using more stories taken from more recent literature and not included in Studies 1, 2, or 3.

Study 4 was also intended to address other concerns. First, it might be argued that confounding variables could explain our finding that health and non-health purity violations are not homogeneous and differ markedly in terms of participant reactions (judgments of emotion, taint to soul, and immorality). Therefore, specifically, we examined variables such as weirdness of the action, harm done to another person, and perpetrator's intention for the action. Previous studies suggest that the weirdness of the actions sampled as purity violations explain the differences between autonomy and purity (Gray & Keeney, 2015). Other studies suggest that

non-health purity violations are more harmful to others than are health violations, which would explain some overlap with autonomy (Dungan et al., 2017; Young & Tsoi, 2013). Other studies highlight the role of intention in judging the morality of any action (Gray et al., 2014; Gray et al., 2012). Therefore, the differences between health and non-health purity violations observed in previous studies might be a function of the differences in weirdness, harm done to another, and intention across these two groups of violations.

Second, we followed-up the question in Study 2 whether participants who identify themselves as politically conservative would judge both health and non-health purity violations more similarly compared to participants who identify themselves as politically liberal. Study 2 found that the purity domain was not homogeneous even for participants who identify themselves as politically conservative: compared to more liberal participants, more conservative participants found non-health purity violations, but not health purity violations, more disgusting and angering. We reexamined this question in Study 4 with a greater number of participants.

Third, in the previous three studies, we argued in favor of thinking of the hypothesized moral domains as being on a continuum rather than being discrete. Our findings showed that anger was the best predictor of immorality: irrespective of the moral domain, the more anger the more immorality. However, arguably, our scatter plots showed health, non-health, and autonomy violations forming three clusters as these violations varied on participant anger, disgust, and grossed-out reactions and immorality judgments. In Study 4, we use autonomy violations pre-selected to be mild with the aim of testing the existence of more overlap between non-health purity violations and autonomy violations.

Finally, to ensure robustness of results across differences in method, there were other changes in details. In contrast to Studies 1 and 2, which had an entirely within-subjects design,

and Study 3, which had both violation type and story as between-subjects factors, Study 4 had violation type nested within stories as a between-subject variable. Similar to previous studies, we included the donating to charity story as a manipulation check.

Details of our materials and analyses plans are pre-registered at <https://doi.org/10.17605/OSF.IO/KC2QS> (Kollareth, 2021).

Method

The method is the same as in Studies 1, 2, and 3 except as noted.

Participants. Participants ($N = 376$) were Americans recruited through Prolific (222 women, 154 men; $M_{\text{age}} = 31.69$ years, range: 18-72 years). An additional 16 participants were dropped for failing to appropriately respond to at least one of the two attention check questions. The attention check consisted of the following question in two different places of the survey: “If you are paying attention to this questionnaire, just click 1 on the following rating scale”.

For a 3 (health vs. non-health vs. harm) x 3 (*anger vs. disgust vs. grossed-out*) Analysis of Variance, our power analysis showed a total sample size of 360 for a medium sized effect ($d = .5$) with the conventional power value of .80 and alpha .05 (two-tailed). A sample size of 360 was also adequate for the multiple regression analysis.

Stories. There were $3 \times 12 = 36$ stories describing actions based on events offered in the literature. There were 24 purity violations (12 physical health and 12 non-health) and 12 autonomy violations. All 36 experimental stories and the source from where they are taken are given in Supplementary Procedures 4.

Unipolar emotion response scales. The emotion response scales were similar to those used in Study 3, but with 5 emotions: *angry, grossed-out, disgusted, sad, and happy*.

Bipolar scale for harm to another. Participants were asked: “Did the person’s actions harm or was it beneficial to another person?” The response format was a 7-point bipolar scale (1-*extremely harming to another*, 2-*moderately harming*, 3-*barely harming*, 4-*neither harming nor beneficial*, 5-*barely beneficial*, 6-*moderately beneficial*, and 7-*extremely beneficial*).

Bipolar scale for intention. Participants were asked: “Was the person’s actions intentional or accidental?” The response format was a 7-point bipolar scale (1-*extremely intentional*, 2-*moderately intentional*, 3-*barely intentional*, 4-*neither intentional nor accidental*, 5-*barely accidental*, 6-*moderately accidental*, 7-*extremely accidental*).

Bipolar scale for weird/atypical. Participants were asked: “Was the person’s actions weird/atypical or normal/typical?” The response format was a 7-point bipolar scale (1-*extremely weird*, 2-*moderately weird*, 3-*barely weird*, 4-*neither weird nor typical*, 5-*barely typical*, 6-*moderately typical*, 7-*extremely typical*).

Design. The design was mixed. There were 12 levels of story nested within of 3 violation types; story was a within-subject factor and violation type was a between-subject factor. Each participant responded to one of the 3 versions of the questionnaire. The order of the 12 stories and the order of dependent variables (taint to the soul scale, emotion scale, morality-immorality scale, harm to another scale, intention scale, and weirdness scale) were randomized separately for each participant.

Procedure. Participants completed the survey online. The participant first read the consent form, then the general instructions, and then the thirteen stories. Participants read the filler story (manipulation check) first and then each of the 12 experimental stories. After reading each story, the participant completed each of the dependent variables. After responding to all 13

stories, participants indicated their political self-identification and answered demographic questions about their sex and age.

Before the analysis, we reverse coded certain scales such that the higher the score the more the taint to the soul, the more harm to another, the more the act was intentional, and the more the act was weird.

Results and Discussion

Check on the bipolar morality-immorality rating scale. Participants' responses to our one moral filler story and autonomy violations indicated that the bipolar immorality rating scale (on which 4 indicated neutral) worked well. Participants judged the action in the moral filler story as *moderately moral* ($M = 2.04$, $SD = 1.34$) and those in the autonomy violation stories, on average, as between *barely* and *moderately immoral* ($M = 5.77$, $SD = 0.59$). (See Supplement S4-1 for ratings for each of the measures for each of the stories).

Emotional responses. Studies 2 and 3 obtained results challenging the theory of a homogeneous and distinct purity domain based on the emotion of disgust. Figure 9 presents mean ratings for each of the three emotions of most interest (*disgust*, *grossed-out*, *anger*) separately for health purity, non-health purity, and autonomy violations based on average scores for 12 stories within each violation type. (Assessment of consistency across the 12 stories within each violation type is reported in Supplement S4-2. Results from the other emotions are presented in Supplement S4-3). The pattern closely matches that described above seen in Studies 2 and 3. The three types of violations did not elicit the same emotional response, as confirmed by a type x emotion interaction, $F(4, 746) = 411.09$, $p < .001$, partial eta squared = .69. Autonomy violations showed the predicted dominant response -- *anger* -- but also a similarly high levels of *disgust*. Collapsed across the health and non-health subtypes, purity violations overall elicited

higher average *disgust* and *grossed-out* than *anger* ratings, as predicted. However, health and non-health purity violations disassociated (as in Studies 1, 2 and 3) based on the relative salience of *anger* and *grossed-out*. Non-health violations were characterized by relatively high levels of *disgust* with lower levels of *grossed-out* and *anger*. In contrast, health violations were characterized by high levels of both *grossed-out* and *disgust*, and very little *anger*.

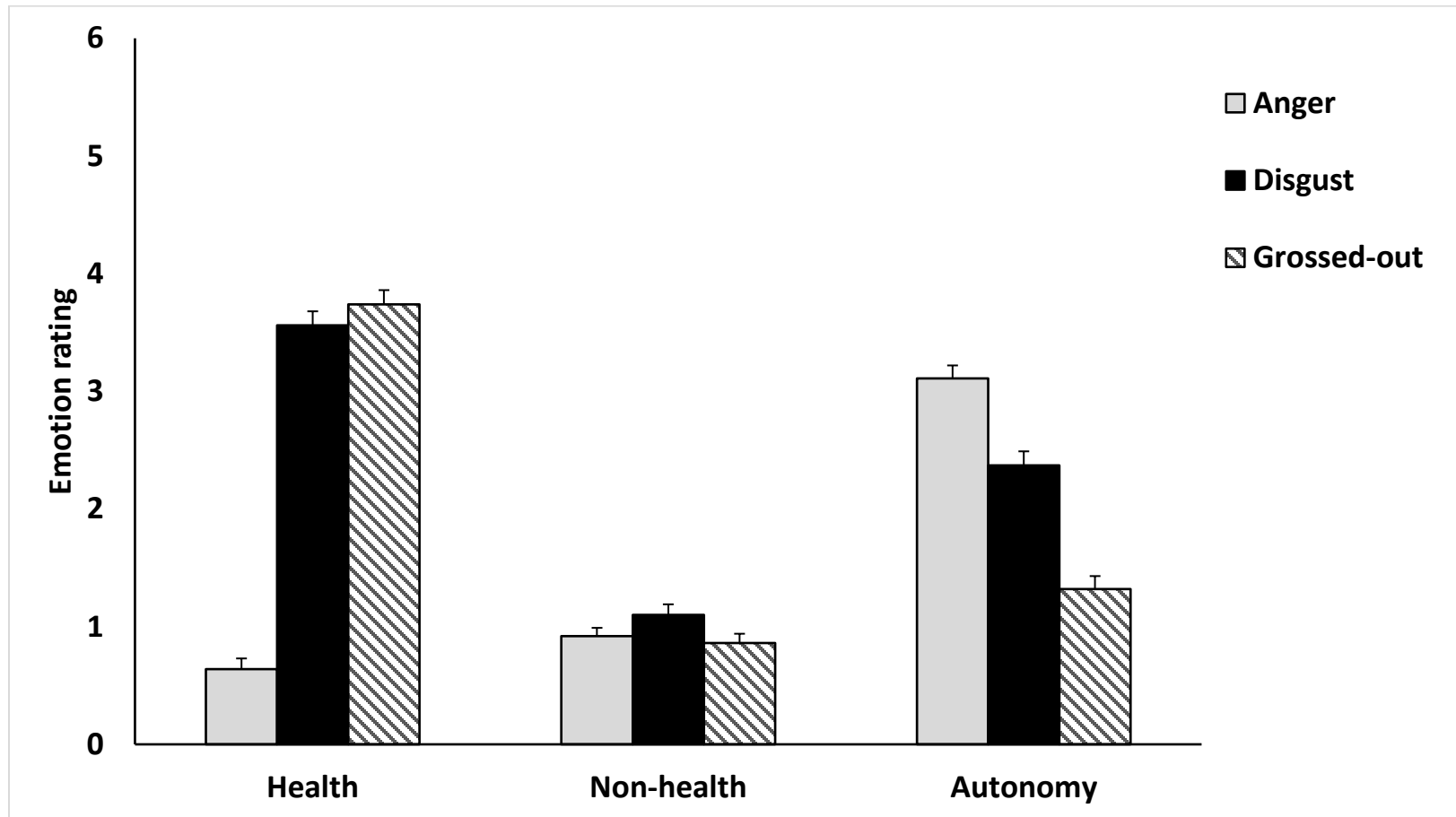


Figure 9. Mean emotion ratings for health, non-health, and autonomy violations, Study 4. Each emotion was rated on an 7-point scale with 0 defined as “no” and 6 as extreme. Standard errors are represented in the figure by the error bars attached to each column.

Follow-up (2 x 2) ANOVAs confirmed the reliability of the interaction pattern in Figure 9. An ANOVA with two violation types (health, non-health) and two levels of emotion (*anger* vs. *disgust*) confirmed a reliable interaction, $F(1, 249) = 447.99, p < .001$, partial eta squared = .64. Similarly, an ANOVA with two violation types (health, non-health) and two of emotion (*disgust* vs. *grossed-out*) also yielded a reliable interaction, $F(1, 249) = 21.95, p < .001$, partial eta squared = .08. For non-health purity stories, due to a floor effect in the data of the present study, we also conducted a sign test, which supported the same conclusion. *Disgust* was higher than *grossed-out* for 71 out of 104 participants with 23 ties excluded, $z = 3.63, p < .001$. Finally, a follow-up ANOVA that compared non-health versus autonomy violations (i.e., excluding health violations) confirmed the predicted switch from *anger* to *disgust* as the primary emotion for non-health violations, $F(1, 250) = 81.58, p < .001$, partial eta squared = .25. These details replicate the relevant effects observed in Studies 1, 2 and 3.

To assess the type x emotion interaction from a different perspective, we analyzed ratings for *anger*, *disgust*, and *grossed-out* separately to evaluate predicted distinctions between purity and autonomy and between health and non-health purity violations. These analyses are reported in Supplement S4-4 and were consistent with the patterns described above.

In sum, witnesses' emotional responses to violations did not point to homogeneity of purity violations or to a sharp distinction between purity and autonomy. Participants disassociated health purity violations from non-health purity violations: Health purity violations elicited mainly *grossed-out* and *disgust* and little *anger*. In contrast, non-health purity violations elicited broadly comparable levels of *disgust*, *grossed-out*, and *anger*. With respect to *anger* and *disgust*, non-health purity violations resembled less severe autonomy violations more than they

did health purity violations. Distinctness was clearly not supported: autonomy violations elicited more feeling of disgust and grossed-out than did non-health purity violations.

Tainting of the soul. Tainting the soul was offered by Rottman et al. (2014) as a marker of a purity violation. Our results replicate findings in Study 3. With our bipolar tainting scale (reverse coded as higher scores showing greater amount of tainting), autonomy violations were judged here as tainting the soul the most—between barely and moderately tainting ($M = 5.64$, $SD = 0.74$); non-health ($M = 4.74$, $SD = 0.56$) and health ($M = 4.50$, $SD = 0.61$) purity violations were between neutral and barely tainting. The main effect of violation type was strong, $F(2, 373) = 109.10$, $p < .001$, partial eta squared = .37. Bonferroni comparisons confirmed that each violation type was significantly different from each of the two other types. Proportion of participants who found the violations as tainting the soul (>4) also showed the same pattern: 83 of 124 participants found health violations tainting the soul to some degree; 112 of 127 participants found non-health violations tainting the soul to some degree; and 113 of 125 participants found the autonomy violations as tainting the soul to some degree. The ordering of violation types is the opposite of that implied by thinking of tainting the soul as a marker of a purity violation; however, note that the order could be altered by varying the severity of infraction across the violation types.

Further analyses question, as in Study 3, the use of taint as a marker of purity violations alone. A multilevel mixed-effects regression model was fit using standardized taint to soul ratings as the outcome variable and was designed to predict variability between participants in outcome from standardized immorality ratings while controlling for differences between stories in outcome from type of violation. There was a strong correlation between how immoral a violation was seen and how tainting to the soul it was considered ($\beta = .72$, $t(30.24) = 32.13$, p

< .001); this correlation was stronger for purity violations combined than for autonomy violations ($t(4159) = 6.36, p < .001$), and stronger for purity non-health and autonomy violations combined compared with purity health ones ($t(3785) = 1.96, p = .051$), although this last difference does not reach the statistical significance threshold. As in Study 1, the more immoral an act, the more it tainted the soul, but in this Study, it was especially the case for purity non-health violations.

Our tentative interpretations are, first, that tainting the soul and immorality were seen as virtually equivalent by our participants, and, second, that stories vary along a single dimension, taint + immorality, that only coincidentally overlaps with violation type. We saw no evidence that taint and immorality judgments are treated differently in different domains.

Morality-immorality. Studies 2 and 3 raised questions about the judged morality-immorality of purity violations. Present results replicated previous results. On average, the autonomy violations were judged the most immoral -- between barely and moderately ($M = 5.77, SD = 0.59$; means for individual stories ranged from 3.00 to 6.75); non-health purity violations were intermediate—between neutral and barely immoral ($M = 4.86, SD = 0.50$; means for individual stories ranged from 3.50 to 6.33); and health purity violations the least immoral-- between neutral and barely immoral ($M = 4.48, SD = 0.67$; means the individual stories ranged from 3.58 to 7.00). See Supplement S4-1 for mean ratings on the morality-immorality scale for each of the stories. For health purity violations, none of the 12 stories were more than barely immoral. The main effect of violation type was strong, $F(2, 373) = 157.58, p < .001$, partial eta squared = .46. Bonferroni comparisons confirmed that each of the three violation types was significantly different from each of the other two.

Controlling for confounding variables. It might be argued that our results that health and non-health differ in participant reactions could be explained by the relative difference in confounding variables such as harm to another, intention of the protagonist, and weirdness. (For more analyses on each of these measures, see Supplement S4-5). Next we examined whether type of violation (health vs. non-health), over and above other variables such as harm to another, intention of the protagonist, and weirdness would predict ratings of *anger*, *disgust*, *grossed-out*, immorality, and taint to soul. In hierarchical multiple regression analysis, emotion ratings (*anger*, *disgust*, and *grossed-out*), morality-immorality rating, and taint-to-soul rating were separately regressed on harm to another, intention of the protagonist, weirdness, and violation type. For each of the dependent variables, a two-step hierarchical multiple regression was conducted with harm to another, intention of the protagonist, and weirdness entered at step 1 and violation type entered at step 2 of the regression. The regression statistics are in Table 2.

Note. Harm to another, intention of the protagonist, and weirdness were entered at step 1 and violation type was added at step 2 of the regression. In addition to harm to another, intention of the protagonist and weirdness, type of violation explained a significant variance in disgust, grossed-out, and immorality. * $p < .05$; ** $p < .001$

Anticipated confounds mattered. Harm done to another person and weirdness of the action consistently predicted ratings for *anger*, *disgust*, *grossed-out*, immorality, and taint to soul (see Table 2). However, they did not fully explain the difference in ratings for the dependent variables across the violation type (health vs. non-health). Type of violation, in addition to harm to another, intention of the protagonist, and weirdness explained 18% of the variance for *disgust* or *grossed-out* ratings: *disgust* ($\Delta R^2 = .183$; $F(1, 246) = 119.15$, $p < .001$); *grossed-out* ($\Delta R^2 = .177$; $F(1, 246) = 139.77$, $p < .001$); 1% of the variance for immorality ratings: ($\Delta R^2 = .014$; $F(1, 246) = 6.28$, $p = .013$). For *anger* and taint to soul ratings, type of violation did not explain any additional variance: *anger* ($\Delta R^2 = .001$; $F(1, 246) = 0.54$, $p = .462$); taint to soul ($\Delta R^2 = .001$; $F(1, 246) = 0.54$, $p = .462$). (See Supplement S4-6 for analysis involving all three violation types).

Political orientation. Our Study 2 showed inconsistent support for the hypothesis that health and non-health purity violations are judged differently by self-identified conservatives than by self-identified liberals. Our findings here partly replicated those from Study 2. As in this previous study, a series of multilevel mixed-effect regression models were used to predict standardized immorality and emotion ratings as the outcome variables, predicting variability between participants in the outcome from standardized responses in political orientation and differences between stories from the type of violation (purity non-health vs. purity health vs. autonomy).

After controlling by the type of violation, immorality ratings were higher as conservatism increased ($\beta = .12$, $t(42.08) = 4.31$, $p < .001$). This correlation was higher for both purity than for autonomy violations ($t(4484) = 4.26$, $p < .001$), but there was no difference between health purity and both autonomy and non-health purity violations combined ($t(4484) = 0.13$, $p = .898$).

Anger responses, on average, were greater for more conservative participants ($\beta = .07$, $t(41.92) = 2.54$, $p = .015$). The correlation between anger and conservatism was stronger for both purity than for autonomy violations ($t(4484) = 2.36$, $p = .019$) but the differences in its magnitude when comparing health purity and both autonomy and non-health purity violations combined were not significant ($t(4484) = 1.86$, $p = .063$).

Disgust responses, on average, were greater for more conservative participants too ($\beta = .08$, $t(57.73) = 2.97$, $p = .004$). The correlation between anger and conservatism was stronger for autonomy and non-health purity violations combined than for health purity violations ($t(4484) = 3.22$, $p = .001$) but no significant differences were found between health purity and both autonomy and non-health purity violations combined ($t(4484) = 1.87$, $p = .062$).

Finally, average grossed-out ratings were not correlated with political orientation ($\beta = .03$, $t(76.62) = 1.15$, $p = .253$). Yet, the magnitude of this correlation was higher for autonomy and non-health purity combined than for health purity violations ($t(4484) = 2.44$, $p = .015$) although there was no difference between both purity and autonomy violations ($t(4484) = 0.52$, $p = .602$).

In sum, as in Study 2, immorality, anger, and disgust ratings increased as the political ideology was more conservative, especially for purity violations and, among them, non-health violations. Grossed-out was not reliably associated with conservatism.

An alternative approach in which political orientation was a dichotomous variable (conservatives, $n = 117$ vs. liberals, $n = 259$) failed to show that the conservatives, compared to liberals, found purity violations as more immoral, *disgusting*, and *gross*. The mean immorality rating was significantly greater for conservatives than for liberals only for non-health violations and not for health and autonomy violations: non-health, conservatives ($n = 39$, $M = 5.06$, $SD =$

0.49), liberals ($n = 88$, $M = 4.76$, $SD = 0.49$), $t(125) = 3.20$, $p = .002$; health, conservatives ($n = 34$, $M = 4.63$, $SD = 0.89$), liberals ($n = 90$, $M = 4.42$, $SD = 0.55$), $t(122) = 1.54$, $p = .126$; autonomy, conservatives ($n = 44$, $M = 5.78$, $SD = 0.65$), liberals ($n = 81$, $M = 5.76$, $SD = 0.56$), $t(123) = 0.21$, $p = .831$.

In a 3 (emotions) x 3 (violation type) x 2 (political orientation) ANOVA, the three way interaction was not significant, $F(4, 740) = 1.24$, $p = .291$, partial eta squared .007. Similar to Study 2, the pattern of emotion ratings across violation types was similar for conservative and liberal participants. This similar pattern is clarified by Figure 10.

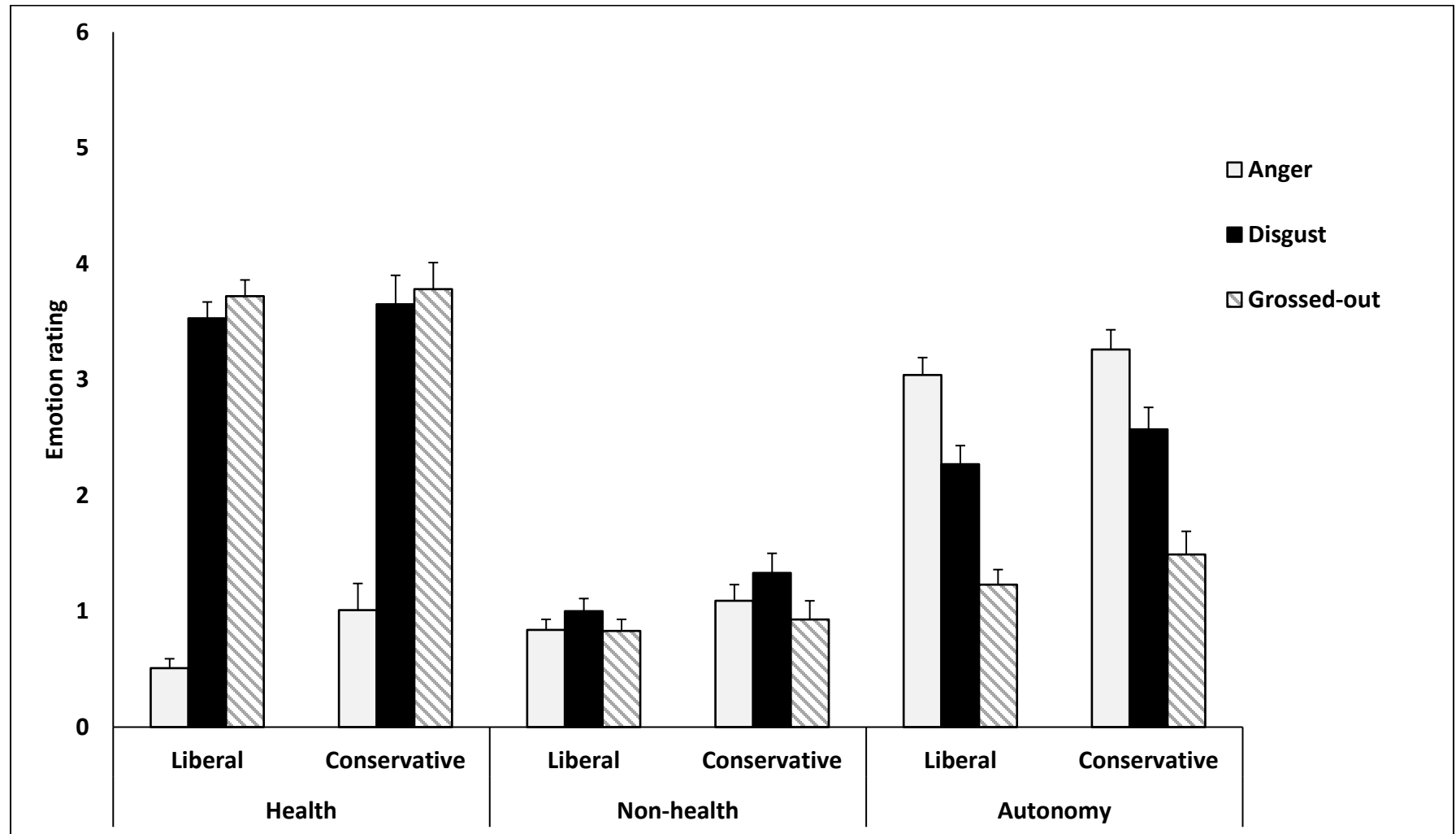


Figure 10. Mean emotion ratings for health, non-health, and autonomy violations for liberal ($n = 259$) and conservative ($n = 117$) participants, Study 4. Standard errors are represented in the figure by the error bars attached to each column.

Similar to Study 2, there was no consistent support for a homogeneous moral domain of purity for conservative participants. In correlational analysis, conservatism was reliably associated with immorality, anger, and disgust, but not grossed-out. MLM analysis showed that both health and non-health violations were more *angering* and immoral for conservative participants than for liberal participants. However, for *disgust*—a hypothesized indicator of purity, correlations within health and non-health found opposing patterns: The more conservative participants found non-health violations more *disgusting*, but health violations less *disgusting*. For *grossed-out*, conservatism showed no reliable association both for health and non-health violations. Our alternative analysis comparing mean average ratings between conservatives and liberals showed that for immorality, conservative participants found non-health violations more immoral than liberal participants. This difference was not found for health violations. The pattern of emotion ratings were similar for conservative and liberal participants.

Continuum across violation-types. Our correlational analyses in Studies 2 and 3 offered evidence challenging the relation of emotion to immorality hypothesized in the theory of a purity domain. Our results showed that purity and autonomy are not distinct domains but may lay on a continuum. The more the *anger*, the more the immorality. In Study 4 we pre-selected autonomy violations to be mild hoping to show more of an overlap among non-health purity and autonomy violations. An additional multilevel mixed-effects regression model was fit using standardized immorality ratings as the outcome variable and was designed to predict variability in outcome between participants from standardized emotional ratings while controlling for the effect of the violation type (purity health vs. purity non-health vs. autonomy violation) in outcome between stories. As in studies 2 and 3, grossed-out ratings were not correlated with immorality ratings ($\beta = .02, t(73.22) = 0.43, p = .672$), and they were dropped from the final

model. Our results are in the predicted direction: According to the final model and in line with the findings in study 2, immorality ratings were higher as both anger and disgust ratings increased ($\beta_{anger} = .36$, $t(65.82) = 9.92$, $p < .001$; $\beta_{disgust} = .37$, $t(39.99) = 9.13$, $p < .001$). The correlation between immorality and disgust ratings was stronger for both purity violations combined than for autonomy violations ($t(4438.52) = 6.21$, $p < .001$) and for non-health purity compared with both health purity and autonomy violations combined ($t(4247.47) = 3.47$, $p = .001$). The correlation between immorality and anger ratings did not change across stories of both purity compared with autonomy violations ($t(4250.51) = 1.89$, $p = .059$), but it was higher for autonomy and purity health violations combined than for purity non-health violations ($t(3618.28) = 2.46$, $p = .014$).

In other words, the more immoral a violation was rated, the more angering and more disgusting it was rated. To illustrate the relation of anger to immorality graphically, Figure 11 shows mean *anger* and mean immorality ratings across the 36 stories. Although inferential tests are not appropriate for these data, the relationship shown was tested from non-aggregated data by MLM.

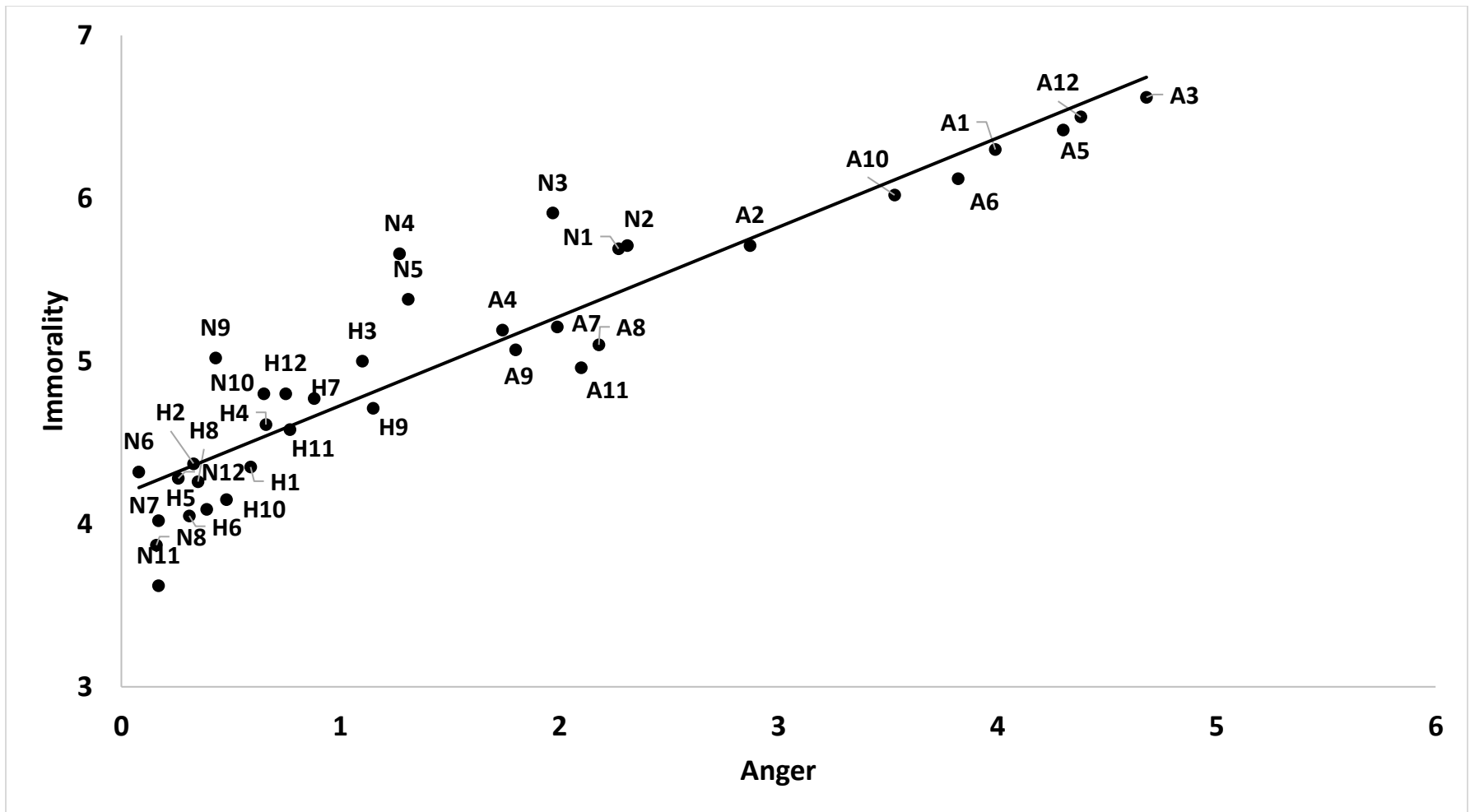


Figure 11. Immorality predicted from anger, Study 4. Both immorality and anger ratings are average ratings for the story (*n* per story ranged from 36 to 66). A1 = Autonomy violation story 1, N1 = Non-health purity violation story 1, and H1 = Health purity violation story 1.

We report more individual difference analyses in Supplement S4-7. Across participants ($N = 376$), immorality was positively and significantly correlated with harm to another ($r(374) = .79, p < .001$), intentionality ($r(374) = .51, p < .001$), and weirdness ($r(374) = .24, p < .001$), further supporting the continuum: the more an act is harmful to another person, intentional, and weird, the more the act is judged immoral. We again followed Rottman et al.'s (2014) suggestion to use correlations to examine individual differences within each violation type.

Testing the homogeneity of health and non-health items by cluster analysis. As outlined under Study 3, we used hierarchical cluster analysis (Between-groups linkage method based on squared Euclidean distance, SPSS Version 27) based on each story's mean *anger*, *disgust*, and *grossed-out* ratings to examine the cohesiveness of the purity and autonomy domains. Results are plotted in Figure 12, in which emotion ratings were used as axes. Cluster analysis revealed three major groupings, one for autonomy items that were high on the anger dimension, another for health items that were low on the anger dimension and high on grossed-out, and, finally, a mixed group of non-health purity and autonomy items that were low on the grossed-out and included considerable spread along the anger dimension. (The dendrogram is presented in Supplement S4-7.) This pattern does not support a single, coherent domain of purity.

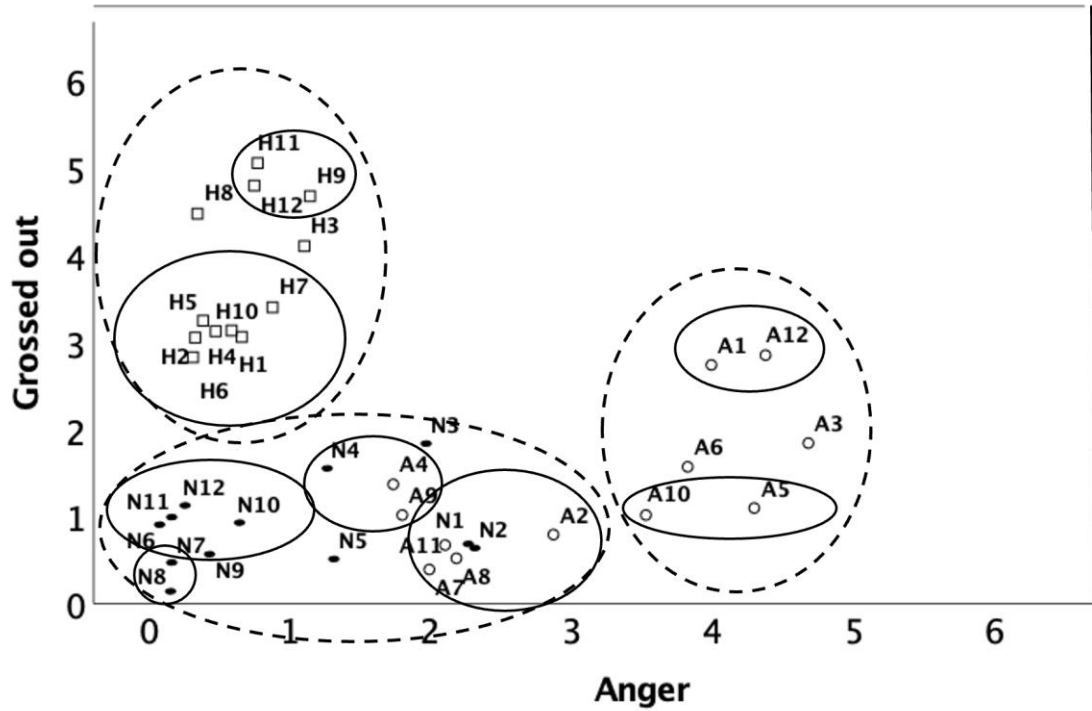


Figure 12. Hierarchical clustering of items based on anger, disgust, and grossed-out ratings. Ratings for each measure within each story is the average rating across the participants (n for health stories = 124; n for non-health stories = 127; n for autonomy stories = 125). A1 = Autonomy violation story 1, N1 = Non-health purity violation story 1, and H1 = Health purity violation story 1. Solid ellipses indicate empirical groups of highly similar items, and dashed ellipses indicate empirical groups of moderately similar stories.

Summary. Moral domain theories hypothesize distinct moral domains, each having homogeneous features. Here we examined whether violations hypothesized as belonging to a purity moral domain would be homogeneously *disgusting* and tainting the soul while distinct from autonomy. We also examined whether the differences observed in previous studies could be explained by the differences in confounding variables.

The emotional profiles observed among health purity, non-health purity, and autonomy violations failed to support traditional domain theory. Instead, they replicated our findings in Studies 2 and 3: For purity, the health violations elicited primarily feeling *grossed-out*, whereas non-health violations elicited primarily *disgust*. For health purity violations, the combination of *grossed-out* and *disgust* was higher than *anger*; whereas, non-health purity violations evoked similar levels of *disgust* and *anger* and, as such, showed greater similarity to participants' reactions to autonomy violations. The autonomy domain elicited the predicted *anger*, but also *disgust* and, to lesser extent, *grossed-out*.

Taint to soul is not a marker of purity violations; our findings here replicated those in Study 3: Violating an autonomy norm was judged more tainting than was violating a purity norm. And within the purity domain, violating a non-health purity norm was more tainting than was violating a health purity norm – which was judged barely tainting if at all. Thus, tainting is a marker, not of purity, but of immorality more generally.

Our Study 4 also confirmed that differences in participant reactions across health versus non-health violations could not be explained by confounding variables, specifically harm to another, intention of the protagonist, and weirdness. Harm to another followed the same pattern as that of immorality and taint to soul: autonomy violations were perceived as harming another person the most, non-health violations intermediate, and health violations minimal. By contrast,

weirdness was similar across the hypothesized domains: Violations within each domain was perceived as between barely and moderately weird. On intentionality, non-health and autonomy violations were perceived as between moderately and extremely intentional whereas health violations, between neutral and barely intentional. More importantly, each of these variables could not fully account for the different pattern of emotional reactions across these violation types. There was also no support for a homogeneous moral domain of purity for conservative participants. The pattern of average emotion ratings were similar for conservative and liberal participants to that of the overall sample.

More generally, Study 4 confirmed the results of Studies 1, 2 and 3: Violations within the hypothesized domain of purity can reliably be subdivided into health and non-health subtypes. In the most relevant dimensions, non-health purity most resembles autonomy.

General Discussion and Conclusion

The key claims of the highly important theory of a distinct moral domain of purity are (a) that it includes threats to both one's physical health as well as to one's soul unrelated to physical health; (b) that both types of threat elicit the same emotion in a witness, an emotion that psychologists call *disgust* (but that lay persons call *grossed-out*); and (c) that violations of purity norms are judged immoral and, uniquely, tainting of the soul. Our findings did not support these hypotheses. Our findings thus suggested a need to reconsider these specific hypotheses and, indeed, the purity domain in general.

Our studies did not address the general theoretical possibility of defining distinct moral domains, only the ones specifically proposed in contemporary moral psychology. Of course, different domains could be defined, even with necessary and sufficient properties: victim is oneself or another, violation occurs on a weekday versus weekend, violation occurs in January,

etc. So, necessary and sufficient defining properties are necessary but not sufficient. The question is whether that property or any cluster of properties helps us understand the psychological basis for the judgment of immorality. We did criticize currently proposed domains for poor definition, but it is the latter point – inability to explain judgments of immorality – that is most telling.

A founding idea in the current theories of discrete moral domains was that each domain is characterized by a common emotional response, one different from the emotion that characterizes any other domain. That emotion in turn was theorized to cause the judgment of immorality (Seidel & Prinz, 2013). The four studies of emotional response reported here found similarities across domains and differences within the purity domain. Autonomy violations elicited the allegedly purity-specific emotion of grossed-out more than did non-health purity violations and elicited disgust more than did either type of purity violation. In short, our results offer a different perspective: Against homogeneous, we say heterogeneous. Against distinct, we say overlapping. Purity and autonomy differ on a continuum. We offer a different possibility about the relationship of these three emotions to moral judgments. Gross is the reaction to health issues, we suggest, whether or not immoral. Rather than causes of a judgement of immorality, anger and disgust are products of that judgment.

The hypothesis suggested by the results here is that non-health purity violations and autonomy violations elicit emotions similar in kind but different in magnitude. On the other hand, purity violations perceived as more immoral could reverse this finding and thus call for a different conclusion. If so, the purity stories examined in the literature confound domain (purity) with mildness of violation. We propose the hypothesis of a continuum: the severity [i.e., how

immoral it is judged to be] of the violation, rather than the domain it supposedly represents, establishes how disgusting, angering, and tainting it is.

More telling were emotional differences within the purity domain. A key methodological step here was including *disgust* and *grossed-out* as separate response options. The emotional response to actions that threaten one's own physical health differed from that to actions that violate alleged purity norms unrelated to health. In four different sets of violations, physical health-related purity violations were judged as principally gross, whereas non-health purity violations were more disgusting than gross (with the exception of sex related norm violations -- "watching pornography," "fantasizing sexual acts"-- as discussed in the Supplement). Our results resonate with previous findings. For example, religious violations free from threats to health do not always elicit the same emotions as to health-related threats (Royzman et al., 2014; Kollareth & Russell, 2019). And *disgust* is associated with immorality beyond purity (Chapman & Anderson, 2014; Steiger & Reyna, 2017). (These findings point to the advantage of including *grossed-out* among the emotion terms studied).

Our supplementary analyses, exploring how *disgust* and *grossed-out* predicted immorality, suggested that the *disgust* rating captures both an *anger* and a *grossed-out* reaction but only the *anger* part of the variance in *disgust* overlaps with immorality. In each of the studies, *grossed-out* was a suppressor variable. The same pattern of results emerged in an analyses across the studies (with stories as cases; $n = 88$). By itself, *disgust* predicted immorality to a moderate degree ($r(86) = .45, p < .001$). By itself, *grossed-out* failed to predict immorality ($r(86) = -.02, p = .883$). Yet, *disgust* and *grossed-out* are intercorrelated ($r(86) = .84, p < .001$). When both *disgust* and *grossed-out* are used to predict immorality in a multiple regression model, R^2 rose to .76, $F(2, 85) = 124.61, p < .001$; with a positive beta weight for *disgust* (beta =

1.59, $p < .001$), but a negative beta weight for *grossed-out*, $\beta = -1.36$, $p < .001$. In the regression equation, the non-overlapping variance in *disgust* can be suppressed: a negative beta weight for *grossed-out* suppresses the *grossed-out* part of *disgust*, leaving the *anger* part of *disgust* to predict immorality. These results are theoretically telling by challenging the claim that judgments of the immorality of purity violations are driven by a basic emotion that psychologists call *disgust* and lay people call *grossed-out*.

Studies 1, 3 and 4 included another measure offered previously as a marker for purity violations— tainting the soul. Again, both similarity across domains and differences within domain were found. Autonomy violations (Studies 3 and 4) were perceived as tainting the soul, indeed, more so than were purity violations. Within the purity domain, health and non-health violations showed different patterns: although neither type of violation was judged as more than moderately tainting and often less, non-health purity violations did show some taint of the soul, whereas health violations showed almost none.

Perceived immorality proved an important factor not sufficiently considered in the past; indeed it was not always even assessed in the past. For exemplars of moral violations, our strategy was to rely on domain-theorist researchers to provide exemplars of violations. But those exemplars were found to differ in perceived immorality in important ways. Autonomy violations were consistently judged to be clearly immoral. Purity violations not so much. What was most surprising, non-health purity violations more generally were judged more immoral than health violations, which were typically judged as barely immoral, if at all. Of the 31 health violations studied here, only one violation (eating one's dog after its natural death) was more than barely immoral, 21 violations were barely immoral, and nine violations were not immoral. Our findings are consistent with a few studies that hint that events that pose a threat to health are judged as

disgusting but not always immoral (Kayyal et al., 2015; Royzman et al., 2009). These results raise the question whether acts previously used in studies of moral violations are all actually perceived as immoral. Undoubtedly, highly immoral purity violations can be found. But then the question arises as to what they elicit and how they differ from the alleged violations included here.

Our inclusion of a bipolar morality-immorality scale raised significant questions about the judged immorality of alleged violations. Future studies in moral psychology would therefore benefit from the use of an explicitly bipolar scale of morality-immorality. More importantly, our surprising results raise a conceptual issue for the theory of purity violations: if some actions deemed by purity theorists as purity violations are moral and other not, then what makes other purity violations immoral?

A critical point is that severity of moral infraction cannot account for our results. Our conclusions are based in large part on interaction effects that reflect different *patterns* of emotion response (feelings of anger, disgust, and grossed-out) across violation types, differences that are independent of main effects of violation type and independent of general effects of severity. Our Study 4 examined autonomy violations that were less extreme, but found a pattern of results were similar to that found in Studies 2 and 3. Because this result might be altered with even less extreme autonomy violations (e.g., making another customer wait in line at a cashier by taking a cell phone call), the relative weighting of individual emotions is more revealing. Of course, in the future, it will be important to examine even more extreme examples of health violations and less extreme autonomy violations.

Our Study 4 also showed that the difference we found between health and non-health violations could not be explained by differences in harm to another person, in the perpetrator's

intention for the action, or in the weirdness of the action. Even after controlling for these variables, health and non-health violations differed in characteristic ways. Another variable of interest was political orientation. Perhaps the purity domain -- both health and non-health violations -- is more evident in politically conservative than in liberal participants. However, our Studies 2 and 4 showed no support for a homogeneous moral domain of purity for just conservative participants. Conservatives, any more than liberals, did not find purity violations (health and non-health) more *disgusting*, *gross* or immoral.

Our studies were limited in scope. As in most studies in this area of research, our measures of emotion were verbal. We had only Americans as participants sampled on the internet. Future studies including more diverse cultural samples might be revealing (Barrett et al., 2016). In other research, we have begun to explore this question (Kollareth et al., 2019). Of course, Americans were the participants in most studies claiming a single purity moral domain. Our studies also used fictitious vignettes and non-ecological samples of moral violations.

The specific sample of purity violations studied is a major limitation. We included 63 different actions (13 in Study 1, 8 in Study 2, 18 in Study 3, 24 in Study 4) characterized as purity violations in previous studies offered as support for the theory of a purity domain. The problem is that it is unclear how exactly to sample exemplars of purity violations. Because the domain of purity is poorly defined, researchers have no definitive way to determine whether or not a specific event is or is not a purity violation. The precise properties that define each domain and that make one domain separate from the others remain elusive, especially for purity. In the present study, we turned therefore to specific acts offered in previous research as violations of purity. It is not clear how those acts were chosen and whether their immorality was an implicit criterion. Some advocates of the theory of a purity domain might object that the 63 actions

included here were poor examples (perhaps insufficiently impure, insufficiently extreme, or insufficiently disgusting), and therefore dismiss our findings.

Some of the 63 violations we studied pose a threat to physical health, and some are unrelated to health, often violating a religious, quasi-religious, or sacred norm. Some actions examined in prior studies included both elements; such confounding rendered results ambiguous on the question whether health and non-health threats belong in the same domain. In the present study, we examined acts previously offered as purity violations but that allowed us to separate those that threaten health from those unrelated to bodily physical health and characterized as posing a threat to one's soul. There were surprisingly few such clear, unconfounded cases.

The problem of sampling raises fundamental issues. Sampling specific acts of purity violation is essential, for comparison between health-related and non-health-related is inherently difficult to interpret because it depends on the specific acts selected. One can imagine choosing acts of either the physical health or non-health type that are more or less outrageous and thereby obtain or fail to obtain a significant difference in the witness's emotional reaction and judgment of immorality. Ideally, the defining properties of the purity domain need be specified, and their ecological distribution taken into account.

We did not include those purity violations from previous studies that confound health-related elements with non-health-related violation of a norm. It might be argued that deconfounding health-related elements from non-health related violation of norms misses the essence of the purity domain. That is, to be a purity violation an action must simultaneously both threaten health and violate a sacred norm. And, indeed, events in which health and sacred norms are combined come readily to mind and perhaps explain the impression that purity is one homogeneous and distinct coherent domain. And use of such events might be what support the

theoretical statements about that domain. Still, scientists would inevitably raise the question whether the witness's emotional reaction of feeling grossed-out stems from the threat to health, whereas the judgement of immorality stems from the sacred norm violation. If so, the question is whether a purity violation so defined is a theoretically interesting class. One could define "weekend sin" as an immoral deed committed on a Saturday or a Sunday, but whether such an arbitrary class is theoretically interesting is doubtful. The question is whether the class of purity violations that threaten both health and sacred norms helps explain their being perceived as immoral.

Our studies suggest that the current concept of a purity norm needs to be reconceptualized. For those who seek to divide morality into domains, one solution consistent with our results would be to replace purity with at least two types of norm. One distinction would be "health-related" purity norms (perceived threats to one's own health or physical wellbeing) versus "non-health" purity norms (perceived threats to one's religious, quasi-religious, or sacred values). "Sex-related" purity norms (perceived threats to one's adherence to standards of sexual propriety) might be a third purity domain. We would then raise our initial questions about each such new domain. How is it defined precisely? Is it homogeneous in the witness's response? Is that response unique to the domain? Are violations of the norm even immoral? Division into more domains, although possible, might not be a winning strategy, and we do not want to be taken as advocating it.

We concede that additional studies are needed to test our conclusions. One approach complementary to that used here would rely on Johnson-Laird and Oatley's linguistic test for semantic hierarchies of emotion terms (Johnson-Laird & Oatley, 1989). We suggest that the word *disgust* is broader than *gross* such that *gross* is a subset of *disgust*. *Gross* implies *disgust*,

but not vice versa. If so, then the sentence “that event is disgusting but not gross” would be acceptable, whereas “that event is gross but not disgusting” would not be.

We also concede a likely historical connection between the two types of violations as we distinguished them. Extending people’s natural response to health threats to violations of norms unrelated to health would support good behavior by attributing negative valence to sacred violations. For example, harsh corporal punishment for a non-health violation (failure to perform a religious rite) would support a practical connection between health and non-health violations. Subsequent threat of both types would elicit fear and avoidance. On the other hand, such connections might be culture-specific and would seem to exist across all domains of morality rather than to be restricted to purity violations. Such connections therefore do not provide any commonality that restricts membership to the purity domain.

Science can be thought of as a dialectic, and we anticipate progress as theses and antitheses are advanced concerning moral purity. Advocates of the current purity domain theory can of course reject our distinctions, or revise them, or provide a refined definition of what is and what is not a violation of the purity domain. Non-advocates can provide an alternative conceptualization of these actions. What is not in dispute is the importance of understanding judgments concerning morality. No more telling judgment exists than that an action is immoral.

References

- Anderson, R. A., & Masicampo, E. J. (2017). Protecting the innocence of youth: Moral sanctity values underlie censorship from young children. *Personality and Social Psychology Bulletin*, 43(11), 1503-1518.
- Arden, M. D., & Rabinovitz, S. (2020). Child sexual abuse and the moralization of purity. *Journal of Child Sexual Abuse*, 29(6), 697-716.
- Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., Gurven, M., . . . Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 113(17), 4688-4693.
doi:<http://dx.doi.org/10.1073/pnas.1522070113>
- Berman, J. Z., & Small, D. A. (2018). Discipline and desire: On the relative importance of willpower and purity in signaling virtue. *Journal of Experimental Social Psychology*, 76, 220-230.
- Berniūnas, R., Silius, V., & Dranseika, V. (2021). Moralization East and West: Moralizing different transgressions among Chinese, Americans, and Lithuanians. *Asian Journal of Social Psychology*. DOI:10.1111/ajsp.12479
- Bucciarelli, M., Khemlani, S., & Johnson-Laird, P. N. (2008). The psychology of moral reasoning. *Judgment and Decision Making*, 3(2), 121-139.
- Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review*, 19(4), 371-394.

- Cannon, P. R., Schnall, S., & White, M. (2011). Transgressions and expressions: Affective facial muscle activity predicts moral judgments. *Social Psychological and Personality Science*, 2(3), 325-331. doi:<http://dx.doi.org/10.1177/1948550610390525>
- Chakroff, A., Dungan, J., Koster-Hale, J., Brown, A., Saxe, R., & Young, L. (2016). When minds matter for moral judgment: intent information is neurally encoded for harmful but not impure acts. *Social Cognitive and Affective Neuroscience*, 11(3), 476-484.
- Chakroff, A., Dungan, J., & Young, L. (2013). Harming ourselves and defiling others: What determines a moral domain?. *PloS one*, 8(9), e74434.
- Chakroff, A., Russell, P. S., Piazza, J., & Young, L. (2017). From impure to harmful: Asymmetric expectations about immoral agents. *Journal of Experimental Social Psychology*, 69, 201-209.
- Chakroff, A., & Young, L. (2015). Harmful situations, impure people: An attribution asymmetry across moral domains. *Cognition*, 136, 30-37.
- Chapman, H. A., & Anderson, A. K. (2014). Trait physical disgust is related to moral judgments outside of the purity domain. *Emotion*, 14(2), 341-348.
- Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior research methods*, 47(4), 1178-1198.
- Dungan, J., Chakroff, A., & Young, L. (2017). The relevance of moral norms in distinct relational contexts: Purity versus harm norms regulate self-directed actions. *PLoS one*, 12(3), e0173405.
- Dungan, J. A., & Young, L. (2019). Asking 'why?' enhances theory of mind when evaluating harm but not purity violations. *Social Cognitive and Affective Neuroscience*, 14(7), 699-708.

- Eibach, R. P., Libby, L. K., & Ehrlinger, J. (2009). Priming family values: How being a parent affects moral evaluations of harmless but offensive acts. *Journal of Experimental Social Psychology, 45*(5), 1160-1163.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.
- Federico, C. M., Ekstrom, P., Tagar, M. R., & Williams, A. L. (2016). Epistemic motivation and the structure of moral intuition: Dispositional need for closure as a predictor of individualizing and binding morality. *European Journal of Personality, 30*(3), 227-239.
doi:<http://dx.doi.org/10.1002/per.2055>
- Franchin, L., Geipel, J., Hadjichristidis, C., & Surian, L. (2019). Many moral buttons or just one? Evidence from emotional facial expressions. *Cognition and Emotion, 33*(5), 943-958.
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47*, 55-130.
- 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.
doi:<http://dx.doi.org/10.1037/a0015141>
- Gray, K. (2014). Harm concerns predict moral judgments of suicide: Comment on Rottman, Kelemen and Young (2014). *Cognition, 133*(1), 329-331.
- Gray, K., DiMaggio, N., Schein, C., & Kachanoff, F. (2021, February 3). What is 'purity'? Conceptual murkiness in moral psychology. <https://doi.org/10.31234/osf.io/vfyut>

- Gray, K., & Keeney, J. E. (2015). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*, 6(8), 859-868.
- 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.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101-124.
- Guerra, V. M., & Giner-Sorolla, R. (2010). The community, autonomy, and divinity scale (CADS): A new tool for the cross-cultural study of morality. *Journal of Cross-Cultural Psychology*, 41(1), 35-50. doi:<http://dx.doi.org/10.1177/0022022109348919>
- Haidt, J. (2003). Elevation and the positive psychology of morality. In C.L.M. Keyes & J. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived* (pp. 275-289). Washington, DC: American Psychological Association.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Pantheon/Random House.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20(1), 98-116. doi:<http://dx.doi.org/10.1007/s11211-007-0034-z>
- Haidt, J., Graham, J., & Joseph, C. (2009). Above and below left-right: Ideological narratives and moral foundations. *Psychological Inquiry*, 20(2-3), 110-119. doi:<http://dx.doi.org/10.1080/10478400903028573>

- Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind, Volume 3: Foundations and the future* (pp. 367-392). New York: Oxford University Press.
- Heerdink, M. W., Koning, L. F., van Doorn, E. A., & Van Kleef, G. A. (2019). Emotions as guardians of group norms: Expressions of anger and disgust drive inferences about autonomy and purity violations. *Cognition and Emotion*, *33*(3), 563-578.
- Herz, R. S., & Hinds, A. (2013). Stealing is not gross: Language distinguishes visceral disgust from moral violations. *The American Journal of Psychology*, *126*, 275-286.
doi:<http://dx.doi.org/10.5406/amerjpsyc.126.3.0275>
- Horberg, E. J., Oveis, C., & Keltner, D. (2011). Emotions as moral amplifiers: An appraisal tendency approach to the influences of distinct emotions upon moral judgment. *Emotion Review*, *3*(3), 237-244.
- Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*(6), 963-976.
doi:<http://dx.doi.org/10.1037/a0017423>
- Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social–functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*(4), 719-737. doi:<http://dx.doi.org/10.1037/a0022408>
- Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and Emotion*, *3*(2), 81-123.

- Kayyal, M. H., Pochedly, J., McCarthy, A., & Russell, J. A. (2015). On the limits of the relation of disgust to judgments of immorality. *Frontiers in Psychology, 6*, 1-9. doi: 10.3389/fpsyg.2015.00951
- Kennedy, B., Atari, M., Davani, A. M., Hoover, J., Omrani, A., Graham, J., & Dehghani, M. (2021). Moral concerns are differentially observable in language. *Cognition, 212*. DOI:10.1016/j.cognition.2021.104696
- Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially purity) help explain culture war attitudes. *Journal of Research in Personality, 46*(2), 184-194. doi:<http://dx.doi.org/10.1016/j.jrp.2012.01.006>
- Koleva, S., Selterman, D., Iyer, R., Ditto, P., & Graham, J. (2014). The moral compass of insecurity: Anxious and avoidant attachment predict moral judgment. *Social Psychological and Personality Science, 5*(2), 185-194. doi:<http://dx.doi.org/10.1177/1948550613490965>
- Kollareth, D. (2021, September 13). Is Purity a Distinct and Homogeneous Domain in Moral Psychology? <https://doi.org/10.17605/OSF.IO/KC2QS>
- Kollareth, D. (2021, November 19). Purity moral domain. Retrieved from osf.io/2rm47
- Kollareth, D., Kikutani, M., Shirai, M., & Russell, J. A. (2019). Do community and autonomy moral violations elicit different emotions? *International Journal of Psychology, 54*(5), 612-620.
- Kollareth, D., & Russell, J. A. (2017). The English word *disgust* has no exact translation in Hindi or Malayalam. *Cognition and Emotion, 31*(6), 1169-1180. doi:<http://dx.doi.org/10.1080/02699931.2016.1202200>
- Kollareth, D., & Russell, J. A. (2019). Disgust and the sacred: Do people react to violations of the sacred with the same emotion they react to something putrid? *Emotion, 19*(1), 37-52.

- Landau, S. I. (Ed.). (2000). *Cambridge Dictionary of American English (Klett Edition) Paperback and CD ROM Pack*. Ernst Klett Sprachen.
- Li, Y., Gao, W., & Chen, M. (2016). Do Chinese distinguish between ethics of community, autonomy and divinity? *Asian Journal of Social Psychology, 19*(1), 26-30.
- Liuzza, M. T., Olofsson, J. K., Cancino-Montecinos, S., & Lindholm, T. (2019). Body odor disgust sensitivity predicts moral harshness toward moral violations of purity. *Frontiers in Psychology, 10*, 458, 1-14.
- Lovett, B. J., Jordan, A. H., & Wiltermuth, S. S. (2012). Individual differences in the moralization of everyday life. *Ethics & Behavior, 22*(4), 248-257.
- Marzillier, S. L., & Davey, G. C. L. (2004). The emotional profiling of disgust-eliciting stimuli: Evidence for primary and complex disgusts. *Cognition and Emotion, 18*, 313-336.
doi:<http://dx.doi.org/10.1080/02699930341000130>
- McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2022). Moral judgment as categorization (MJAC). *Perspectives on Psychological Science, 17*(1), 131-152.
- Milesi, P. (2016). Moral foundations and political attitudes: The moderating role of political sophistication. *International Journal of Psychology, 51*(4), 252-260.
doi:<http://dx.doi.org/10.1002/ijop.12158>
- Murray, D. R., & Schaller, M. (2016). The behavioral immune system: Implications for social cognition, social interaction, and social influence. In J. M. Olson, & M. P. Zanna (Eds.), *Advances in experimental social psychology* (pp. 75-129, Chapter vii, 323 Pages). San Diego, CA: Elsevier Academic Press.

- Nabi, R. L. (2002). The theoretical versus the lay meaning of disgust: Implications for emotion research. *Cognition and Emotion*, *16*(5), 695-703.
doi:<http://dx.doi.org/10.1080/02699930143000437>
- Piazza, J., Sousa, P., Rottman, J., & Syropoulos, S. (2019). Which appraisals are foundational to moral judgment? Harm, injustice, and beyond. *Social Psychological and Personality Science*, *10*(7), 903-913.
- Pilotti, M. A., & El Alaoui, K. F. (2019). A study of memory, evaluation, and choice with an (un) clean conscience. *Australian Journal of Psychology*, *71*(3), 203-211.
- Rieger, N. S., Worley, N. B., Ng, A. J., Christianson, J. P. (2022). Insular cortex modulates social avoidance of sick rats. *Behavioural Brain Research*, *416*, 113541.
- Ritter, R. S., Preston, J. L., Salomon, E., & Relihan-Johnson, D. (2016). Imagine no religion: Heretical disgust, anger and the symbolic purity of mind. *Cognition and Emotion*, *30*(4), 778-796.
- Rottman, J., Kelemen, D., & Young, L. (2014). Tainting the soul: Purity concerns predict moral judgments of suicide. *Cognition*, *130*(2), 217-226.
doi:<http://dx.doi.org/10.1016/j.cognition.2013.11.007>
- Rottman, J., & Young, L. (2019). Specks of dirt and tons of pain: Dosage distinguishes impurity from harm. *Psychological Science*, *30*(8), 1151-1160.
- Royzman, E. B., 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. doi:<http://dx.doi.org/10.1037/a0036829>
- Royzman, E., & Kurzban, R. (2011). Minding the metaphor: The elusive character of moral disgust. *Emotion Review*, *3*(3), 269-271. Doi: <http://dx.doi.org/10.1016/j.tics.2013.06.001>

- Royzman, E. B., Leeman, R. F., & Baron, J. (2009). Unsentimental ethics: Towards a content-specific account of the moral-conventional distinction. *Cognition*, *112*(1), 159-174. doi: <http://dx.doi.org/10.1016/j.cognition.2009.04.004>
- Rozin, P., Haidt, J., & Fincher, K. (2009). From oral to moral. *Science*, *323*(5918), 1179-1180.
- Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 757-776). New York, NY: Guilford Press.
- 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. doi: <http://dx.doi.org/10.1037/0022-3514.76.4.574>
- Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125*(1), 3-30. doi: <http://dx.doi.org/10.1037/0033-2909.125.1.3>
- Schaller, M. (2016). The behavioral immune system. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (2nd ed. pp. 206-224, Chapter xxvi, 660 Pages). Hoboken, NJ: John Wiley & Sons Inc.
- Schaller, M., & Park, J. H. (2011). The behavioral immune system (and why it matters). *Current Directions in Psychological Science*, *20*(2), 99-103.
doi:<http://dx.doi.org/10.1177/0963721411402596>
- 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.

- Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, 22(1), 32-70.
- Schein, C., Ritter, R. S., & Gray, K. (2016). Harm mediates the disgust-immorality link. *Emotion*, 16(6), 862-876.
- Scott, S. E., Inbar, Y., & Rozin, P. (2016). Evidence for absolute moral opposition to genetically modified food in the United States. *Perspectives on Psychological Science*, 11(3), 315-324.
- Seidel, A., & Prinz, J. (2013). Sound morality: Irritating and icky noises amplify judgments in divergent moral domains. *Cognition*, 127(1), 1-5.
doi:<http://dx.doi.org/10.1016/j.cognition.2012.11.004>
- Simpson, A., & Rios, K. (2017). The moral contents of anti-atheist prejudice (and why atheists should care about it). *European Journal of Social Psychology*, 47(4), 501-508.
doi:<http://dx.doi.org/10.1002/ejsp.2219>
- Steiger, R. L., & Reyna, C. (2017). Trait contempt, anger, disgust, and moral foundation values. *Personality and Individual Differences*, 113, 125-135.
- Sweetman, J., & Newman, G. A. (2020). Replicating different roles of intent across moral domains. *Royal Society Open Science*, 7: 190808. <http://dx.doi.org/10.1098/rsos.190808>
- Tracy, J. L., Steckler, C. M., & Heltzel, G. (2019). The physiological basis of psychological disgust and moral judgments. *Journal of Personality and Social Psychology*, 116(1), 15-32.
- Vasquez, K., Keltner, D., Ebenbach, D. H., & Banaszynski, T. L. (2001). Cultural variation and similarity in moral rhetorics: Voices from the Philippines and the United States. *Journal of Cross-Cultural Psychology*, 32(1), 93-120.

Vezzali, L., Brambilla, M., Giovannini, D., & Paolo Colucci, F. (2017). Strengthening purity:

Moral purity as a mediator of direct and extended cross-group friendships on sexual prejudice. *Journal of Homosexuality*, 64(6), 716-730.

Young, L., & Tsoi, L. (2013). When mental states matter, when they don't, and what that means

for morality. *Social and Personality Psychology Compass*, 7(8), 585-604.