

Harmful situations, impure people: an attribution asymmetry across moral domains

Alek Chakroff¹

Liane Young²

1. Harvard University, Cambridge, MA

2. Boston College, Chestnut Hill, MA

Word Count: 5734

Corresponding Author:
Alek Chakroff
33 Kirkland St.
Cambridge, MA 02138
chakroff@fas.harvard.edu

Abstract

People make inferences about the actions of others, assessing whether an act is best explained by person-based versus situation-based accounts. Here we examine people's explanations for **norm violations** in different domains: harmful acts (e.g., assault) and impure acts (e.g., incest). Across four studies, we find evidence for an attribution asymmetry: people endorse more person-based attributions for impure versus harmful acts. This attribution asymmetry is partly explained by the abnormality of impure versus harmful acts, but not by differences in the moral wrongness or the statistical frequency of these acts. Finally, this asymmetry persists even when the situational factors that lead an agent to act impurely are stipulated. These results suggest that, relative to harmful acts, impure acts are linked to person-based attributions.

Keywords: morality; moral psychology; moral foundation theory; attribution theory; action explanation; social cognition

1. Introduction.

Upon hearing about the **abhorrent** actions of others, such as the torture of inmates by Americans in Abu Ghraib, many people demand an explanation: why did they do it? People generate spontaneous explanations of others' behavior, which also support evaluations of the offending agents and their actions. Do people attribute the cruel acts of Abu Ghraib to the stable dispositions of the torturers, or to their extreme situational pressures? Research in social psychology indicates that people often underestimate the impact of situational influences when assessing the behavior of others (e.g., actor-observer asymmetry; Gilbert & Mallone, 1995; Jones & Harris, 1967; Jones & Nisbett, 1971; Malle, Knobe, & Nelson, 2007; Milgram, 1963; Nisbett, Caputo, Legant, & Marecek, 1973; Ross, 1977). Here we investigate when and why people attribute **norm-violating behavior** to the person versus the situation, taking the contrast between harmful acts (e.g., assault) and disgusting acts (e.g., pathogen exposure) as a case study.

Prior work has indicated significant variability in people's explanations of others' behavior. For example, people assign greater causal and intentional responsibility to those who commit morally wrong versus neutral acts (Alicke, 1992, 2000; Knobe, 2006; Leslie, Knobe, & Cohen, 2006). Some people pass moral judgment on many different kinds of action, from harmful acts that negatively affect others, to victimless (yet odd or disgusting) acts that people perform in the privacy of their own homes. Do people's action explanations depend on the kind of violation they are judging? Recent work suggests that people are sensitive to different sources of information when making moral judgments of actions that are primarily *angering / harmful* (e.g., assault) versus actions that are primarily *disgusting / impure* (e.g., eating taboo foods). For example, when condemning impure acts versus harmful acts, people are less sensitive to the context of the act and the mental state of the actor (Chakroff, Dungan, & Young, 2013; Russell

& Giner-Sorolla, 2011a; Young & Saxe, 2011), as well as any potential reasons or mitigating circumstances in explaining the impure action (Rozin, Imada, Lowery, & Haidt, 1999; Russell & Giner-Sorolla, 2011b; Russell, Piazza, Giner-Sorolla, 2013). These findings support theories that there are different kinds or domains of morality (Graham et al., 2011; Haidt, 2007). Prior work has also highlighted individual differences in moral judgments across moral domains – in particular, impure acts are judged as more morally wrong by social conservatives versus liberals (e.g., Haidt, 2007). However, even among liberals, moral judgments of impure acts (e.g., suicide) are not at floor and are best predicted by purity concerns (Rottman, Kelemen, & Young, 2014). Notably, in the present work, moral judgments of impure acts were above floor in all studies (all t 's versus floor were > 11 , p 's $< .001$, see Table 2 for means and SEs).

The present research tests the following hypothesis: people are more likely to explain *impure* acts using *person-based attributions*, appealing to factors internal to the person who committed the violation, but people are more likely to explain *harmful* acts using *situation-based attributions*, appealing to factors external to the person who committed the violation. This hypothesis follows from prior work indicating a dissociation between moral judgments that focus on the moral permissibility of an action and moral judgments that focus on the actor's moral character (Chakroff et al., 2013; Tannenbaum, Uhlmann, & Diermeier, 2011; Uhlmann & Zhu, 2014). For example, participants judge “beating your girlfriend” to be a morally worse *action* than “beating your girlfriend's cat”, but they judge the cat-beater to be a morally worse *person* (i.e., to have poorer moral character) compared to the girlfriend-beater (Tannenbaum et al., 2011). Another study revealed harsher moral judgments of *impure* acts when participants focused on the moral status of the *person* but harsher judgments of *harmful* acts when participants focused on the moral status of the *action* (Chakroff et al., 2013). Here, we hypothesized that, in addition to

person-based *moral judgments*, impure acts would be associated with increased person-based *attributions* as well.

Across four studies, we examine the effect of moral domain (i.e., harm versus impurity) on action explanations. Study 1 tests the hypothesis that people explain impure versus harmful acts using more person-based versus situation-based attributions and indicates (as with all studies) that the effect cannot be explained by domain differences in moral wrongness. Study 2 explores a potential mediator of the attribution asymmetry across domains: the perceived abnormality of impure versus harmful acts. Study 3 replicates the mediation findings from Study 2 and indicates further that the attribution asymmetry cannot be accounted for by judgments of the frequency (e.g., how commonly an act is performed) of impure versus harmful acts. Finally, Study 4 tests whether the domain attribution asymmetry persists even when the situational influences that led to the violations are stipulated within the scenarios (i.e., agents are *forced* to perform certain impure or harmful action).

1.2. Methodological Notes.

For all studies, participants were recruited using Amazon Mechanical Turk (www.mturk.com). All participants completed an IRB-approved online consenting process. Participants were self-reported English speakers and were excluded from analyses only if they did not complete the survey, or if they completed the survey in one minute or less, i.e., a conservative estimate of the minimum amount of time participants could spend thoughtfully reading and completing the survey (consent, scenarios, and demographics). The stimuli used in Studies 1-3 are reported in Table 1, while the full stimuli and measures for all studies are reported in the Supplementary Material. All studies were analyzed using linear mixed effects models, with item and participant as random factors, and domain and target (when applicable) as

fixed factors. All results hold when modeling effects on condition averages using ANOVAs (see Supplementary Material).

2. Study 1: Person-based explanations for impure acts

Study 1 tests the hypothesis that people attribute impure acts more to an agent's disposition, and less to the situation, compared to harmful acts. We tested this hypothesis using a survey adapted from measures used in classic work on the actor-observer asymmetry (Nisbett et al., 1973). However, rather than asking participants to provide trait ascriptions (e.g., I am "energetic," "relaxed," or "depends on the situation"), we asked participants to indicate whether a person (either oneself or another person) was the type to engage in (or never engage in) harmful or impure behaviors, or whether a person might engage in the target behaviors depending on the situation.

2.1. Method.

2.1.1. Participants. 258 American participants (65% female; $M_{age} = 34.7$) were recruited in an online sample and were paid \$0.41 for their time. Sixty-two participants were excluded according to the exclusion criteria above.

2.1.2. Procedure. Participants read 9 items describing harmful behavior and 9 items describing impure behavior (see Table 1)¹; items were presented in random order. In a between-participants manipulation, action explanations were made either for oneself or for another person ("the average person"). For each item, participants were forced to choose among three response options: "I am (the average person is) the type to do this," "I am (the average person is) the type

¹ An exploratory factor analysis indicated that each subset of 9 items formed a coherent cluster, with all factor loadings $> .71$ for harmful items within the first factor, and all factor loadings $> .82$ for impure items within the second factor, based on judgments of anger and disgust (Study 3). Cronbach's α for harmful and impure items = .94 and .97, respectively.

not to do this," or "There are situations that could lead me (the average person) to do this," where the first two choices represent person-based attributions for a behavior, while the latter choice represents situation-based attributions (Cf., Nisbett et al., 1973). On a separate page, participants then rated "how morally wrong" each item was on a 5-point scale from "not at all wrong" to "extremely wrong".

<i>Domain</i>	
<i>Harmful</i>	<i>Impure</i>
Kick someone in the shin	Lick someone's shoe
Scratch someone on the arm	Pour urine on oneself
Prick someone's hand with a needle	Pick up dog poop barehanded
Poke someone in the eye	Step in vomit barefoot
Burn someone on the arm	Taste earwax
Hit someone's finger with a hammer	Drink cow blood
Cut someone's cheek with a razor	Pick up a snot-filled Kleenex
Pinch someone hard on the arm	Pick up a used Band-Aid
Whip someone with a belt	Eat a worm

Table 1. Harmful and impure acts used in Studies 1-3.

2.2. Results and Discussion.

2.2.1. Effects of domain and target on measures. Condition means and standard errors are reported in Table 2, with attribution results also displayed in Figure 1. For attributions, we report the mean proportion of person-based attributions, with "1" indicating fully person-based attributions across items and participants, and "0" indicating fully situation-based attributions. Participants endorsed more person-based attributions for impure versus harmful acts ($F(1, 2674.2) = 90.96, p < .001$)². Collapsing across domains, participants also made more person-based attributions for their own acts ($M = .841; SE = .019$) relative to others' acts ($M = .695; SE$

² The present results remain significant when analyzing the "is the type not to do this" responses alone, see Supplementary Information.

= .019) ($F(1, 189.9) = 29.12, p < .001$), in contrast to past work on the actor-observer asymmetry (Nisbett et al., 1973). Finally, in a domain x target interaction ($F(1, 2674.2) = 29.71, p < .001$), the domain attribution asymmetry was stronger for judgments of “the average person” (impure $M = .79, SE = .02$; harmful $M = .60, SE = .02$; $F(1, 1510.8) = 95.81, p < .001$) relative to judgments about oneself (impure $M = .87, SE = .02$; harmful $M = .82, SE = .02$; $F(1, 1174.4) = 10.05, p = .002$).

2.2.2. Relationships between measures. Correlational analyses revealed that morally worse acts were linked to more person-based attributions ($r(196) = .31, p < .001$). However, moral judgments were also harsher for harmful versus impure acts ($F(1, 2629.6) = 1172.98, p < .001$), suggesting that the greater proportion of person-based attributions for impure versus harmful acts cannot be accounted for by differences in moral judgment alone. Indeed, when entering moral judgments into the linear mixed effects model, the effect of domain on attributions remains significant ($F(1, 2663.2) = 92.88, p < .001$).

3. Study 2: Person-based explanations and Abnormality

Study 1 indicates that people are more likely to endorse person-based versus situation-based attributions for impure acts compared to harmful acts. However, the underlying source of the asymmetry remains open for investigation. Study 2 examines the role of a potential mediating variable, the abnormality of the act, in the domain attribution asymmetry found in Study 1. Abnormal acts diverge from the norm by definition and break convention. Breaking mere conventional norms, such as wearing pajamas to work, is not necessarily morally wrong (e.g., Kelley et al., 2007; Turiel, 1979). However, abnormal acts may still be informative of a person’s character (e.g., Tannenbaum et al., 2011) and therefore drive person-based attributions.

3.1. Method.

3.1.1. Participants. 105 American participants (61% female; $M_{age} = 30.1$) were recruited in an online sample and were paid \$0.36 for their time. Five participants were excluded.

3.1.2. Procedure. Participants read the same 18 items from Study 1, presented in random order. Participants read all items three times in separate blocks, delivering judgments of (1) action explanation, (2) moral wrongness, and (3) abnormality. Block order was randomized across participants.

For attributions, participants endorsed either a person-based or situation-based attribution for each behavior. Unlike Study 1, participants made judgments of “the average person” and chose between two options: "The average person is either the type to do this, or to never do this" (person) or "There are situations that could lead the average person to do this" (situation). Participants delivered moral judgments as in Study 1. For abnormality judgments, participants rated “How weird (i.e., unusual, bizarre, odd)” each item was on a 5-point scale from "not at all weird" to "extremely weird."

3.2. Results and Discussion.

3.2.1. Effect of domain on measures. Condition means and standard errors are reported in Table 2, with attribution results also displayed in Figure 1. As in Study 1, participants endorsed more person-based attributions for impure versus harmful acts ($F(1, 1501.6) = 236.10, p < .001$), and harsher moral judgments for harmful versus impure acts ($F(1, 1433.3) = 664.83, p < .001$). In line with our predictions, participants delivered greater abnormality judgments for impure versus harmful acts ($F(1, 1523.5) = 520.42, p < .001$).

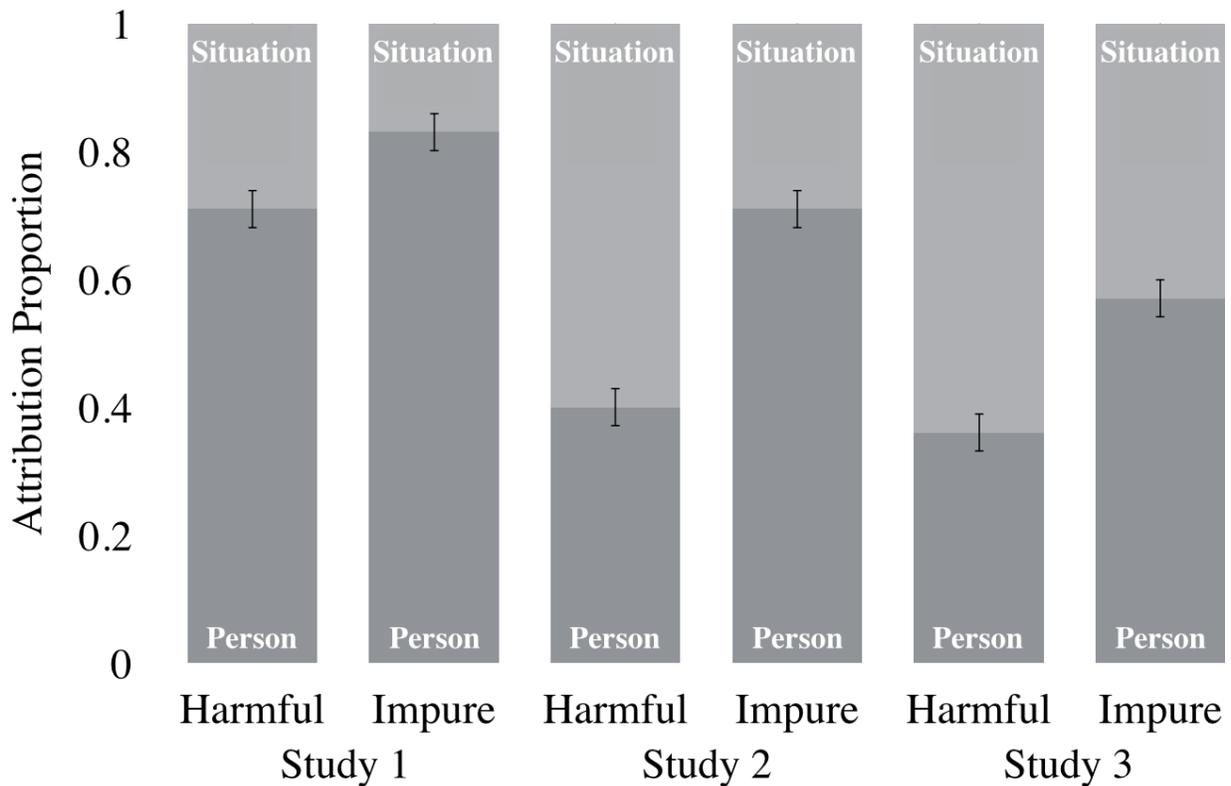


Figure 1. Attributions for others' harmful and impure acts. Error bars indicate +/- 1 SEM.

3.2.2. *Relationships between measures.* Averaging across all stimuli, correlations revealed that, as in Study 1, people endorsed more person-based attributions for morally worse acts ($r(100) = .33, p = .001$). People also endorsed more person-based attributions for more abnormal acts ($r(100) = .51, p < .001$). However, a linear regression indicated that attributions were uniquely predicted by abnormality judgments ($\beta = -0.46, p < .001$) but not moral judgments ($\beta = -0.09, p = .4$). We then examined the effect of domain on attribution via abnormality and moral wrongness using multiple mediation analysis on stimuli means within each domain (e.g., Preacher & Hayes, 2008). When simultaneously entering moral wrongness and abnormality as potential mediators, the effect of domain on attributions was mediated by both moral wrongness and abnormality, but in opposing directions. Participants judged impure acts as less morally

wrong than harmful acts, which would predict more person-based attributions for harmful than impure acts (indirect effect = -0.08, Sobel $Z = -3.25$, $p = .001$); however, impure acts were judged as more abnormal than harmful acts, leading to more person-based attributions for impure than harmful acts (indirect effect = .11, Sobel $Z = 4.50$, $p < .001$). Despite significant mediation, domain remained a direct predictor of action explanations in the mediation model (direct effect = .18, $p < .001$). This pattern of data suggests that, as in Study 1, impure acts elicited more person-based explanations, *despite* being judged as less morally wrong than harmful acts. However, the domain attribution asymmetry remained significant even when statistically controlling for the abnormality of the acts via mediation analysis.

		Harmful	Impure
Study 1	Person Attribution	0.71 (0.02)	0.83 (0.02)
	Morally Wrong	4.10 (0.06)	2.80 (0.05)
Study 2	Person Attribution	0.40 (0.03)	0.71 (0.03)
	Morally Wrong	3.50 (0.09)	2.20 (0.08)
	Abnormal	3.00 (0.08)	4.10 (0.08)
Study 3	Person Attribution	0.36 (0.04)	0.57 (0.04)
	Moral Action	3.74 (0.08)	2.60 (0.09)
	Moral Character	3.52 (0.08)	3.48 (0.08)
	Abnormal	3.06 (0.07)	4.15 (0.07)
	Frequent	2.07 (0.07)	1.41 (0.06)
	Angering	3.90 (0.08)	2.26 (0.08)
	Disgusting	2.51 (0.08)	4.25 (0.07)

Table 2. Ratings in Studies 1-3 (means, standard errors). Person attribution ratings are proportions of person-based attributions (out of 1) averaged across participants, reported separately when averaging across harmful and impure stimuli. All other ratings are made on 5-point Likert scales with anchors “Not at all X” (1) and “Extremely X” (5), where X was a measure-relevant descriptor (e.g., “Not at all disgusting”).

4. Study 3: Person-based explanations, Abnormality, and Frequency

Study 2 indicated that the domain attribution asymmetry was partially mediated by the perceived abnormality of impure versus harmful acts. Study 3 explores two senses in which

abnormality may be conceptualized (e.g., prescriptive vs. statistical; Hitchcock & Knobe, 2009).

The first sense, which we call prescriptive *abnormality*, concerns the extent to which an act departs from norms or breaks conventional rules. The second sense refers to the statistical *frequency* of an action in everyday life. The two kinds of abnormality are asymmetrically related, such that norm-violating actions are relatively infrequent or atypical, but infrequent or atypical acts are not necessarily norm-violating. For example, the total fertility rate in America is less than 2 lifetime births per woman (World Bank), making the act of giving birth an exceedingly infrequent act. Births to loving, married couples are even less frequent but certainly do not violate norms in most people's eyes.

Theories of action explanation suggest that the relative frequency of an act can influence people's explanations of that act (Kelley, 1967). For example, if Jane sees a movie and thinks it is hilarious, but no one else does, it may be reasonable to explain Jane's amusement by appealing to her disposition (e.g., she is easily amused), rather than the movie itself (e.g., it's a comedy). Study 3 examines whether it is the perceived frequency of an act that is predictive of person-based attributions in the context of harmful and impure moral violations.

In Study 3, we also examine two sets of additional measures, derived from two factors that distinguish between impure and harmful acts. The first factor denotes a *moral focus* on the *action* versus the *person* who committed the act. In prior work, moral judgments of harms were harsher when participants focused on the *action*, while moral judgments of purity violations were harsher when participants focused on the agent's *character* (Chakroff et al., 2013). The second factor of interest was *emotion*, e.g., anger and disgust. Mounting research suggests that harmful and impure acts are judged as eliciting relatively more anger and disgust, respectively (Russell & Giner-Sorolla, 2011b). We investigated both factors to determine whether the attribution

asymmetry for impure versus harmful acts could be explained by domain differences in appraisals of anger and disgust or moral judgments focusing on the action or the agent.

4.1. Method.

4.1.1. Participants. 114 American participants (60% female; $M_{age} = 35.6$) were recruited in an online sample and were paid \$0.41 for their time. Thirteen participants were excluded.

4.1.2. Procedure. Participants read the same 18 items from Studies 1-2 describing harmful and impure behaviors, presented in random order. Participants read all items eight times in separate blocks, delivering judgments of (1-2) action explanation, (3) moral action, (4) moral character, (5) abnormality, (6), frequency, (7) disgust, and (8) anger. Block order was randomized across participants.

For action explanations, participants endorsed either a person-based or situation-based attribution for each behavior. While in Studies 1-2, participants made judgments about “the average person,” in Study 3 participants simply delivered attributions for “a person,” allaying potential concerns that participants in Study 2 were conveying the notion that “the average person” is unlikely to perform abnormal acts. Participants also made more explicit causal attributions in response to a novel measure, endorsing statements that a behavior was either “More caused by the person” or “More caused by the situation.”

Unlike Studies 1-2, participants delivered two kinds of moral judgments designed to focus on the moral status of the act itself (moral action) (e.g., How morally wrong is this behavior?), versus judgments that focus on the moral character of the agent (moral character) (e.g., How “sick” or “screwed up” is this person?) (Chakroff et al., 2013; Tannenbaum et al., 2011). Abnormality judgments were identical to those in Study 2. For frequency judgments,

participants judged “Is this behavior typical or common in the general population?”. For the disgust measure, participants judged “Is this behavior gross?”. For the anger measure, participants judged “Is this behavior angering?”. With the exception of action explanations, all ratings were made on 5-point Likert scales with anchors “Not at all X” (1) and “Extremely X” (5), where X was a measure-relevant descriptor (e.g., “Not at all gross”).

4.2. Results and Discussion.

4.2.1. Effect of domain on measures. Condition means and standard errors are reported in Table 2, with attribution results also displayed in Figure 1. Both attribution measures were simultaneously entered into a linear mixed effects model, with across-measure variance captured in a random factor. As in Studies 1 and 2, participants endorsed more person-based explanations for impure versus harmful acts ($F(1, 1433.3) = 664.83, p < .001$). Notably, the novel, explicit attribution measure captured the domain attribution effect in isolation ($F(1, 1564.2) = 149.30, p < .001$). Also, as in Studies 1 and 2, participants delivered harsher moral judgments for harmful versus impure acts ($F(1, 1574.8) = 516.63, p < .001$). Judgments of moral character did not differ for harmful and impure acts ($F(1, 1595.9) = 0.34, p = .56$). Nonetheless, as in prior work (Chakroff et al., 2013), impure acts were judged as worse when participants focused on the moral character of the agent versus the moral status of the action, while harmful acts were judged as worse when participants focused on the action versus the agent’s character, reflecting a domain x moral focus interaction ($F(1, 1850) = 240.04, p < .001$).

As in Study 2, participants delivered greater abnormality judgments for impure versus harmful acts ($F(1, 1532.3) = 454.66, p < .001$). Participants also judged impure acts as being less frequent, compared to harmful acts ($F(1, 1314.1) = 332.10, p < .001$). Finally, participants judged harmful acts as more angering than impure acts ($F(1, 1662.8) = 1100.34, p < .001$), and impure

acts as more disgusting than harmful acts ($F(1,1278.4) = 1141.49, p < .001$) (Cf. Russell & Giner-Sorolla, 2011b).

4.2.2. Relationships between measures. We averaged across both attribution measures ($r(202) = .44, p < .001$) to form a composite measure of person-based attribution. Separate correlations revealed that attribution was best predicted by the abnormality of an act ($r(202) = .47, p < .001$), followed by moral judgments focusing on the agent's character ($r(202) = .43, p < .001$), judgments of disgust ($r(202) = .35, p < .001$), and judgments of the frequency of the act ($r(202) = -.15, p = .03$). Attributions were not significantly predicted by moral judgments focusing on the action ($r(202) = .08, p = .26$) or judgments of anger ($r(202) = .02, p = .76$). Entering abnormality, character, disgust, and frequency judgments into a linear regression indicated that attributions were uniquely predicted by both abnormality judgments ($\beta = .30, p = .001$) and moral judgments of character ($\beta = .26, p < .001$).

As in Study 2, we averaged across stimuli within each domain and examined the effect of domain on attribution via our measures using multiple mediation analysis (e.g., Preacher & Hayes, 2008). Unlike Study 2, only abnormality remained a significant mediator of domain on attribution (indirect effect = 0.07, Sobel $Z = 2.72, p = .007$). Also unlike Study 2, the mediation was full, with no significant direct effect of domain on attribution (see Fig 2A). A follow-up analysis testing only abnormality as mediator suggested that the full mediation was driven by abnormality alone (indirect effect = 0.11, Sobel $Z = 4.75, p < .001$), and not an additive effect of the other measures (see Fig. 2B).

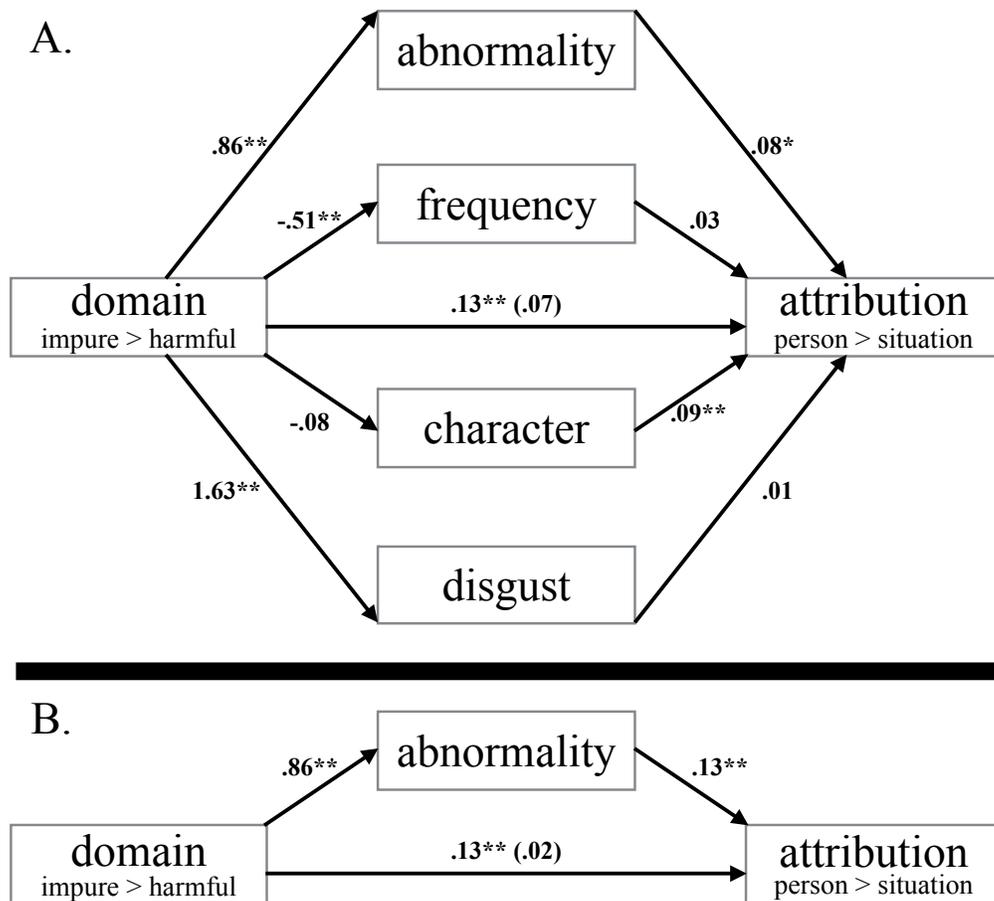


Figure 2. Study 2, Standardized regression coefficients for relationship between domain and attributions, as mediated via (A) abnormality, frequency, character, and disgust, and via (B) abnormality alone. ** $p < .001$, * $p < .01$, else $p > .1$.

As expected, moral character was a powerful predictor of attribution. However, unlike prior work (Chakroff et al., 2013), character judgments did not differ across domains, leaving little room for the measure to explain the domain effect on attribution. Critically, although frequency judgments significantly differed across domains and predicted attribution in isolation of other measures, frequency did not remain a significant unique predictor of attribution alongside abnormality. Indeed, partial correlations reveal that abnormality remains predictive of

attribution when controlling for frequency ($r(199) = .46, p < .001$), while frequency does not when controlling for abnormality ($r(199) = .07, p = .32$).

5. Study 4: Force and Freedom

Studies 1-3 indicate that people are more likely to endorse person-based versus situation-based explanations for impure acts compared to harmful acts. However, it may be reasonable for participants to respond as they did: there may be fewer situations that could lead one to perform an impure act (e.g., drinking cow's blood) versus a harmful act (e.g., whipping someone with a belt). Do person-based explanations for impure acts persist even when the situational constraints that led an agent to perform a harmful versus impure act are stipulated? In Study 4, participants are informed of how situational pressures lead an agent to act in harmful or impure ways. In each scenario, an agent is forced (by an external agent) to commit a specific moral violation (e.g., incest versus assault).

Study 4 also differs from Studies 1-3 in the measures used. Instead of collecting action explanations in binary categories (person versus situation), we borrow measures from Young & Phillips (2011) to examine the degree to which participants judge an agent to have been forced to act versus to have freely acted. These judgments indicate the extent that participants perceive the agent to be the primary cause of the moral violation, acting freely and not forced by the situation.

5.1. Method.

5.1.1. Participants. 281 participants³ (39% female; $M_{age} = 30.2$) were recruited in an online sample and were paid \$0.11 cents for their time. Forty-four participants were excluded.

5.1.2. Procedure. Participants read scenarios involving two agents, one of whom (the agent) was forced by the other (the forcer) to commit a moral violation. Each scenario was

³ Location data were not collected for this sample.

chosen from a pair of thematically matched stories that differed in the kind of moral domain involved (i.e., harm versus impurity). Participants read two scenarios involving violations from the same domain.

For example, participants read a scenario about a game of truth or dare, where a player, John (the agent), is forced by another player (the forcer) to either punch his sister (harm) or kiss his sister (impurity). Participants then made judgments of force (e.g., Was John forced to do this?), freedom (e.g., Was John free to do otherwise?), blame assigned to the agent (e.g., Should John be blamed for his actions?), blame assigned to the forcer (e.g., Should the other player be blamed for his actions?), and the general moral status of the scenario (e.g., Was this situation morally upsetting?). Each question was presented on a separate page; question order was randomized, with the exception of force, which always appeared first, and moral wrongness, which always appeared last.

5.2. Results and Discussion.

5.2.1. Effect of domain on measures. Judgments of force (e.g., Was John forced to do x? *reverse-scored*) and freedom (e.g., Was John free to do otherwise?) ($r(237) = .48, p < .001$) were combined into a composite measure of agent-freedom for the remaining analyses. As expected, a main effect of domain on agent-freedom ($F(1, 231) = 4.42, p = .037$) indicated that agents' acts were judged as more freely performed (and *less* forced) when the acts were impure ($M = 5.28; SE = .13$) versus harmful ($M = 4.90; SE = .13$), consistent with the domain attribution asymmetry reported in Studies 1 through 3. There was a significant effect of domain on agent-blame ($F(1, 244.8) = 6.46, p = .012$), with more blame assigned to harmful agents ($M = 5.60; SE = 0.13$) versus impure agents ($M = 5.14; SE = 0.13$). There was no significant effect on forcer-blame ($F(1, 246.4) = 2.35, p = .13$), or general moral status ($F(1, 244.8) = 0.1, p = .76$),

suggesting more generally that domain differences in action explanation are not driven by domain differences in moral wrongness.

5.2.2. Relationships between measures. Correlations revealed relationships between greater agent-freedom and greater agent-blame ($r(237) = .29, p < .001$) and between greater agent-freedom and reduced forcer-blame ($r(237) = -.29, p < .001$), but no relationship between agent-freedom and general moral status ($r(237) = .02, p = .73$). Follow-up analyses indicated that the main effect of domain on agent-freedom persisted when controlling for agent-blame ($F(1, 227.1) = 7.59, p = .006$), forcer-blame ($F(1, 227.6) = 4.30, p = .039$), or general moral status ($F(1, 230.8) = 4.30, p = .039$).

6. General Discussion

Across four studies, we find evidence for **an** attribution asymmetry: people explain impure (versus harmful) acts using dispositional, person-based attributions, and people explain harmful (versus impure) acts using situational attributions. The effect was not explained by domain differences in moral wrongness (Studies 1-4) but was tied to the perceived abnormality of impure versus harmful acts (Studies 2-3). Notably, the effect persisted even when the situational factors leading to harmful or impure acts were specified within the stimuli (Study 4).

The present work supports the idea that there are at least two distinct senses of abnormality (Hitchcock & Knobe, 2009). First, abnormality may reflect the degree to which an act violates local cultural norms. Second, abnormality may reflect the statistical infrequency or atypicality of an act. The covariation model of attribution (Kelley, 1967) stipulates that (in)frequency should predict person-based attributions. Indeed, zero order correlations suggest that this is the case ($r(202) = -.15, p = .03$). However, abnormality judgments were more strongly

related to attributions ($r(202) = .47, p < .001$), and this relationship accounted for variance in frequency judgments in predicting attributions.

What do judgments of abnormality reflect, if not (in)frequency? Both harmful and impure acts violate norms, but impure acts were judged as more abnormal than harms (Studies 2 & 3). Participants are not providing binary judgments of whether an act violates norms or not, but rather they are delivering quantitative judgments of the degree that the act violates norms of behavior. Participants may be evaluating abnormality judgments within a framework of action prediction and expectation (Hitchcock & Knobe, 2009). People expect others to conform to prescriptive norms. To illustrate, suppose John is boarding a bus. Knowing nothing about John, can we estimate the likelihood that he will (1) Pull out cash and pay to get on, (2) Try to get on without paying, and (3) Pull out cash and eat it in front of the driver? Both 2 and 3 are unexpected and violate norms of behavior, but we see 3 as both less expected and a greater violation of norms of behavior. Critically, the degree of norm violation is not the same as the moral wrongness of a violation - we see 2 as being more morally wrong, but 3 as violating norms to a greater degree within an action prediction framework. Notably, harmful acts were judged as more morally wrong than impure acts in the present work, while impure acts were judged as more abnormal than harmful acts, underscoring this distinction.

People may view abnormal violators as having a deviant or warped personality. Notably, Study 3 also measured moral judgments of an agent's character (Cf., Chakroff et al., 2013; Tannenbaum et al., 2011; Uhlmann & Zhu, 2014) as a candidate mediator of the domain effect. Although character judgments were strongly predictive of attributions ($r(202) = .43, p < .001$), they did not differ across harmful and impure agents ($t(200) = 0.58, p = .56$; see Table 2), in contrast to prior work (Chakroff et al., 2013). This leaves little room for moral character, as

operationalized in the present work, to explain the attribution asymmetry. However, abnormality judgments may reflect judgments of some aspects of abnormal character. In other words, not only is the behavior seen as “weird,” so too is the person.

The present results add to a literature suggesting that people judge harmful and impure acts to be immoral via separable cognitive processes. Past work has shown that judgments of impure acts are less sensitive to contextual factors, including information about the actor’s intentions or reasons – e.g., unintentionally impure acts are still wrong (Chakroff et al., 2013; Russell & Giner-Sorolla, 2011, 2013; Young & Saxe, 2011). Based on this past literature alongside the current findings, we suggest that impure acts are specifically informative about the disposition of the *person* who engaged in the act, both in terms of the person’s moral character (Chakroff et al., 2013) and in terms of the person’s causal role in generating the behavior in the first place.

These results also buttress recent research exploring people’s moral judgments of self-directed harm (Chakroff et al., 2013; Rottman et al., 2014). In one study, participants judged others who commit suicide (versus homicide) to have a “tainted soul”, and these judgments are predictive of judgments of suicide as morally wrong overall (Rottman et al., 2014). In another study, moral judgments of self-directed harmful and impure acts were judged as indicating the poor moral character of the offending agent (Chakroff et al., 2013). The present work suggests that, in addition to having a poor moral character and a tainted soul, impure others may also be judged as more personally causally responsible for their actions, regardless of the circumstances. Future research is needed to directly connect person-based moral judgments, action explanations, and perceptions of a tainted soul.

The present work may shed light not only on lay perceptions of suicide and self-harm but also on lay perceptions of sexual assault and a tragic consequence of thereof: judgment of the sexual assault victim as somehow blameworthy or causally responsible for the event, in spite of the force of the perpetrator (Bieneck & Krahe, 2011; Niemi & Young, in prep). The present results indicate that those who are forced to perform impure acts are judged to have acted more freely, compared to those who are forced into harmful action (Study 3). Perhaps, in the case of impure actions, people are less capable of distinguishing between the **offending** agent and the victim, both when delivering moral judgments and action explanations.

Because the domain attribution asymmetry is positively correlated with the abnormality of the act (Study 2), the effect may also increase as a function of the abnormality of the *situation* that leads to the act. As such, people may attribute the actions of individuals in extremely abnormal situations to dispositional qualities of the individuals. This may explain why people's "attribution errors" and consequent shock at other people's behavior so often occur when judging those in bizarre situations (Gilbert & Mallone, 1995; Jones & Harris, 1967; Milgram, 1963; Ross, 1977; Zimbardo, 1973). Consider the events following the crash of Flight 571 in the Andes in 1972. Those who survived resorted to cannibalism in order to stay alive (Parrado & Rause, 2007). Although these individuals did not directly harm anyone, as they merely ate the flesh of deceased passengers, our data suggest that people may judge the survivors to be personally causally responsible for their actions, and perhaps to have tainted moral character as well (Chakroff et al., 2013). Understanding the cognitive processes underlying the attribution asymmetry across domains is critical to interpreting our inferences about others' actions.

Acknowledgments

Many thanks to Jon Baron, Amelia Brown, James Dungan, Brendan Gaesser, Kurt Gray, Larisa Heiphetz, Laura Niemi, Jared Piazza, Josh Rottman, Jordan Theriault, and Lily Tsoi for comments and advice. This project was made possible by a grant from the John Templeton Foundation and a grant from the Alfred P. Sloan Foundation.

References

- Alicke, M. D. (1992). Culpable causation. *Journal of personality and social psychology*, 63(3), 368-378.
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological bulletin*, 126(4), 556.
- Bieneck, S., & Krahe, B. (2011). Blaming the Victim and Exonerating the Perpetrator in Cases of Rape and Robbery: Is There a Double Standard?. *Journal of interpersonal violence*, 26(9), 1785-1797.
- Chakroff, A., Dungan, J., Young, L. (2013). Harming ourselves and defiling others: what determines a moral domain? *PLoS ONE*, 8(9).
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological bulletin*, 117, 21-21.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of personality and social psychology*, 101(2), 366.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101-124.
- Gray, K., & Schein, C. (2012). Two minds vs. two philosophies: Mind perception defines morality and dissolves the debate between deontology and utilitarianism. *Review of Philosophy and Psychology*, 3(3), 405-423.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2), 389-400.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998-1002.
- Hitchcock, C. & Knobe, J. (2009). Cause and Norm. *Journal of Philosophy* 11, 587-612.
- Inbar, Y., Pizarro, D. A., & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition and Emotion*, 23(4), 714-725.
- Inbar, Y., Pizarro, D., Iyer, R., & Haidt, J. (2012). Disgust sensitivity, political conservatism, and voting. *Social Psychological and Personality Science*, 3(5), 537-544.
- Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of experimental social psychology*, 3(1), 1-24.
- Jones, E. E., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior* (p. 16). New York: General Learning Press.
- Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language*, 22(2), 117-131.
- Kelley, H. H. (1967). Attribution theory in social psychology. In *Nebraska symposium on motivation*. University of Nebraska Press.

- Knobe, J. (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies*, 130(2), 203-231.
- Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting Intentionally and the Side-Effect Effect Theory of Mind and Moral Judgment. *Psychological Science*, 17(5), 421-427.
- Malle, B. F., Knobe, J. M., & Nelson, S. E. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology*, 93(4), 491.
- Milgram, S. (1963). Behavioral Study in Obedience. *Journal of Abnormal Psychology*, 67, 371-378.
- Niemi, L. & Young, L. (in preparation). Contamination and condemnation: Moral judgment in the case of sexual assault.
- Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology*, 27(2), 154-164.
- Parrado, N., & Rause, V. (2007). *Miracle in the Andes: 72 days on the mountain and my long trek home*. Three Rivers Press.
- Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. *The social psychology of morality: Exploring the causes of good and evil*, 91-108.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879-891.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in experimental social psychology*, 10, 173-220.
- 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.
- Rottman, J., Kelemen, D., & Young, L. (2014). Tainting the soul: Purity concerns predict moral judgments of suicide. *Cognition*, 130(2), 217-26.
- Russell, P. S., & Giner-Sorolla, R. (2011a). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, 11(2), 233-240.
- Russell, P. S., & Giner-Sorolla, R. (2011b). Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science* 2(4), 360-364.
- Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: What it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological bulletin*, 139(2), 328.
- Russell, P. S., Piazza, J., & Giner-Sorolla, R. (2013). CAD Revisited Effects of the Word Moral on the Moral Relevance of Disgust (and Other Emotions). *Social Psychological and Personality Science*, 4(1), 62-68.
- Schein, C., & Gray, K. (2014). The Prototype Model of Blame: Freeing Moral Cognition From Linearity and Little Boxes. *Psychological Inquiry*, 25(2), 236-240.
- Tannenbaum, D., Uhlmann, E. L., & Diermeier, D. (2011). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology*, 47(6), 1249-1254.
- Uhlmann, E.L., & Zhu, L. (2014). Acts, persons, and intuitions: Person-centered cues and gut reactions to harmless transgressions. *Social Psychological and Personality Science*, 5, 279-285.

- Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition, 120*(2), 202-214.
- Turiel, E. 1979 : Distinct conceptual and developmental domains: social convention and morality. In Howe, H. and Keasey, C. (eds), *Nebraska Symposium on Motivation, 1977: Social Cognitive Development*. Lincoln : University of Nebraska Press.
- Zimbardo, P. G. (1973). On the ethics of intervention in human psychological research: With special reference to the Stanford Prison Experiment. *Cognition, 2*(2), 243–256.

[Click here to download Supplementary Item: ChakroffYoung_SM.docx](#)

Highlights:

- Across four studies, participants explain the impure versus harmful acts of others using person-based attributions.
- The attribution effect is not explained by differences in moral wrongness across domains.
- The attribution effect is explained in part by differences in perceptions of the abnormality of acts across domains, but not by the statistical frequency or typicality of the acts.
- The attribution effect persists even when the situational causes of harmful and impure acts are stipulated.