

**Linguistic evidence for the dissociation between impurity and harm:
Differences in the duration and scope of contamination versus injury**

Laura Niemi, PhD*

Cristina Leone, BSc

Liane Young, PhD

Corresponding author:
Laura Niemi, niemi@cornell.edu
Cornell University
Department of Psychology

DISSOCIATING IMPURITY AND HARM

Abstract

Previous research has shown that harm and impurity are relevant to a different extent across individuals and transgressions. However, the source of these differences is still unclear. Here, we combine language analysis and social-moral psychology to articulate the core defining features of impurity versus harm. In Study 1(a-c), we found systematic variation in language use indicating that people infer that contamination, unlike injury, affects a target completely and irreversibly, rendering them a transmitter of contamination. In Study 2(a-b), we investigated how evoking intuitions about these core features of contamination—completeness, irreversibility, and transferability—influences judgments of impurity and harm. We found that implying effects on a target were complete and irreversible altered judgments of impurity, but not harm. Overall, our research supports the conclusion that impurity and harm are substantially distinct in cognition and moral judgment: unlike harm, impurity connotes negative effects that spread continually across space and time.

Keywords: moral psychology; language analysis; harm; impurity; contamination; blame

Introduction

Consider a nation facing a crisis, where two groups of citizens have joined together to oust the leaders. One group spreads the message that the leadership *harms* the nation's standing in the eyes of the world; another group spreads the message that the leadership *taints* the nation's standing in the eyes of the world. Such a minor tweak to language might seem inconsequential. Indeed, some research in moral psychology has argued that judgments of impurity are just a functionally equivalent subset of harm judgments. Other research, however, has shown that people direct harm and impurity judgments at different kinds of transgressions, and to a differing extent across individuals. While harm and impurity both communicate that a target is strongly negatively affected, they are invoked differently, for reasons that have not yet been clarified by existing research. The present research takes a novel approach to understanding the nature and moral implications of harm and impurity judgments. Here, we investigate the core defining features of injury and contamination, how these features are reflected in language, and how they influence moral judgments.

The current research builds on a growing literature in moral psychology concerned with understanding how people delineate morality. For example, recent work has addressed fairness versus loyalty, "binding" versus "individualizing" values (Graham et al., 2011, 2018; Haidt, 2007; Hildreth, Gino & Bazerman, 2016; Mooijman et al., 2018; Napier & Luguri, 2013; Niemi & Young, 2013; 2016; Niemi, Hartshorne, Gerstenberg, Stanley & Young, in press; Niemi, Wasserman & Young, 2018; Waytz, Dungan & Young, 2013), as well as harm versus impurity judgments (Parkinson & Byrne, 2018; Dungan, Chakroff & Young, 2017; Chakroff, Dungan & Young, 2013; Chakroff & Young, 2015; Chakroff et al., 2016; Gray, Schein & Ward, 2014; Young & Saxe, 2011). Some researchers have argued that impure acts are simply instances of

DISSOCIATING IMPURITY AND HARM

harm, referring to work indicating that participants seem to conceptually prioritize harm over impurity, refer to the harmful nature of impure acts, and identify “victims” of ostensibly victimless purity violations (Gray et al., 2014; Schein, Ritter & Gray, 2016; Schein & Gray, 2018). Other researchers of Moral Foundations Theory, looking to the possible evolutionary roots of these moral norms, contend that the domains of harm and impurity are disassociated: harm norms protect *individuals* regardless of group membership, whereas purity norms protect *groups* by obligating individuals to adhere to collective norms (Haidt, 2003; 2007; Graham et al., 2011). Dungan and colleagues (2017) likewise argue that impurity judgments do not collapse into harm judgments: they find that impure, unlike harmful, acts are judged as most immoral when committed against *oneself* rather than others. Other work has distinguished harm and impurity by examining associated cognitive patterns. People are more likely to explain impure versus harmful acts with reference to the person rather than the situation (Chakroff & Young, 2015); and, moral judgments of harm are modulated by information about agents’ intentions, whereas moral judgments of impurity are relatively insensitive to information about intent (Chakroff et al., 2016; Barrett et al., 2016; Chakroff et al., 2013; Young & Saxe, 2011). Emotion researchers have investigated impurity and harm by comparing the cognition associated with disgust and anger-inducing acts, respectively. They have found, for example, that participants employ simplistic reasoning to justify moral condemnation of disgust-eliciting sexual acts versus anger-eliciting harmful acts—reflected in greater use of evaluative (e.g., “X is bad”) versus elaborative judgments (e.g., “X violated others’ human rights”), respectively (Russell & Giner-Sorolla, 2011). Convergent evidence reveals disgust to be generally resistant to reasoning. For example, participants report a distaste for juice that has come in contact with a sterilized

DISSOCIATING IMPURITY AND HARM

cockroach and feces-shaped fudge, even when both are perfectly safe to consume (Rozin, Millman, Nemeroff, 1986).

The prior research has largely used stimulus sets of transgressions that ostensibly represent acts of harm and impurity. Disagreement about the particular characteristics (e.g., weirdness, Gray & Keeney, 2015; dose sensitivity, Rottman & Young, 2019) that qualify a transgression as harm versus impurity has prevented researchers from coming to a consensus about the distinctiveness of these moral domains. In the current research, we bypass these disagreements with an approach that does not rely on stipulating behaviors as either harmful or impure. Instead, we investigate these domains by studying the structure and use of moral language, thereby discovering (rather than stipulating) what constitutes harm or impurity.

Research on moral language has the potential to advance understanding as it merges theory-driven and data-driven approaches: natural language serves as a window into cognition, as theories of language and moral psychology guide the interpretation of linguistic data. For example, some prior research on moral language has investigated the moral discourse characterizing social networks and debates around pressing social issues; this research has applied Moral Foundations Theory to interpret patterns in the use of words referring to the moral foundations (e.g., harm, fairness, loyalty, purity; Dehghani et al., 2016; Clifford & Jerit, 2013; Sagi & Dehghani, 2014). This exercise has revealed that arguments on topics such as stem cell research, the World Trade Center attacks, and abortion, use words related to the moral foundations differently; notably, differences often reflect variation in references to purity and harm. This research, and much of the prior work on moral language, has helped to shed light on the features of moral discourse.

DISSOCIATING IMPURITY AND HARM

Here, we pursue a different but complementary language-based approach. We build on theories of social-moral cognition and cognitive linguistics, proposing that impurity and harm meaningfully dissociate: language used to communicate about impurity, but not harm, encodes specific information about the duration and scope of negative effects on a target — namely, the completeness, irreversibility, and transferability of the damage. This not to say that impurity and harm have nothing in common; they both entail negatively affecting a target. However, a complete understanding of impurity and harm requires considering this shared feature in the context of their important differences.

A prominent account of moral cognition, the theory of dyadic morality (Gray, Waytz & Young, 2012; Schein & Gray, 2018), is consistent with the shared nature of impurity and harm. On this account, the framework of a moral transgression involves an agent who is causal and intentional, harming a patient who is not. This framework informs moral judgment, motivating evaluations such as blame and wrongness directed at the agent. Squaring the theory of dyadic morality with the findings of other researchers of harm and impurity is problematic. Impurity can be imparted without the existence of a dyad, such as when one finds oneself impure through one's own actions (Dungan et al., 2017); and, people harshly judge *unintentional* acts that impart impurity (e.g., Chakroff et al., 2016; Young & Saxe, 2011). Furthermore, as detailed in the next section, linguistic theory on which the dyadic morality theory is built does not support overlapping psychological frameworks for impurity and harm.

Other social psychological research has identified common principles of thinking that are good fits with the concept of impurity in a moral context. The “law of contagion” is a pattern of thought that researchers have found U. S. adults apply in thinking about morally relevant “contagions” (Rozin et al., 1986; Nemeroff & Rozin, 1994). The law of contagion stipulates that

DISSOCIATING IMPURITY AND HARM

an essence is transferred by contact between entities (no intention necessary), and continues to affect the contacted entity long after contact ends, sometimes forever. People’s use of the contagion template is consistent with the features of contamination posited to set impurity apart from harm in the present research: negative effects that are complete, irreversible, and transfer to others.

In the current studies, we combine the methods of linguistic analysis and social psychology vignette studies to investigate the defining features of harm and impurity. As described in the next section, we analyze the meaning and usage of words for harm and impurity, revealing essential differences in how these processes transmit their negative effects (Study 1 a-c). Subsequently, in Study 2 (a-b), we investigate how the core features of impurity – complete and irreversible negative effects, that transfer to others – influence contamination and harm judgments.

Contamination and Injury in Language

The semantics of verbs that communicate impurity, such as *contaminate* and *taint*, contrast in important ways with verbs that convey harm, such as *injure* and *wound*. According to a prominent theory of verb semantics (Levin, 1993; Kipper, Korhonen, Ryan & Palmer, 2008; Kipper-Schuler, 2005), *contaminate* and *taint* are classified as members of the *fill* category (9.8). These verbs have semantic frameworks that indicate that a substance moved with the result of an entity that is completely affected. For example, when “A tainted B” this can

<p>FILL 9.8 <i>adorn, anoint, bathe, coat, contaminate, cover, soil, sully, surround, swathe, taint,¹</i></p>

¹ Sampling of the 97 verbs in the “fill” class; See page 20 of the Supplementary Material for full list.

DISSOCIATING IMPURITY AND HARM

indicate an event in which an agent (A) causes a substance to relocate somewhere new (B), or, an event in which some substance (A) and its destination (B) become co-located².

Verbs in the *fill* category, including *contaminate* and *taint*, have a particular grammatical feature; they do not allow “holistic/partitive” alternation (Levin, 1993, p. 50; VerbNet, Kipper-Schuler, 2005), unlike a range of other verbs such as *spray* which do allow this alternation, and can be used in both “partitive” or “holistic” constructions. For example:

(1) Holistic construction, “*I sprayed the sink with water*”

(2) Partitive construction: “*I sprayed water into the sink*”

The meaning of these sentences may differ, but both types are felicitous – appropriately constructed and smoothly understood. The holistic construction (1) implies that the target (the sink) was completely affected whereas the partitive construction (2) allows for the interpretation that some parts of the target (sink) may have been unaffected by the water (Levin, 1993).

The non-alternating verbs *contaminate* and *taint*, however, accommodate only the holistic construction and not the partitive construction. For example:

(3) Holistic construction: “*He tainted the woman with evil*”

(4) Partitive construction: “*He tainted evil into the woman*”

The felicitous (3) holistic construction implies that target (the woman) was completely affected by the tainting. The infelicitous (4) partitive construction forces an interpretation that is incongruent with the semantics of *taint*: namely, that some parts of the target may have been unaffected by the tainting. The fact that verbs that convey impurity are felicitous with the holistic

² For example, the felicitous sentence “He tainted their souls” fits with framework (A) which does not specify which substance that the agent, “He”, caused to be relocated to the destination, “their souls.” The felicitous sentence “Evil tainted their souls” fits with the agent-less framework (B) where the substance, “evil,” is co-located with the destination, “their souls.”

DISSOCIATING IMPURITY AND HARM

construction, but infelicitous with the partitive construction, supports the hypothesis that people expect contamination and taint to involve a process that affects an entity completely.

By contrast, the verbs *injure* and *wound* do not necessarily evoke “complete affectedness” of the object, nor do many other verbs in harm events like *hit* and *stab*. *Injure* is among the “*hurt*” verbs (Levin, 1993, p. 34) which can take as their direct objects particular *parts* of the body (e.g., “*She injured her hand*”). This feature is consistent with the harm’s reversibility: unlike impurity and contamination, injuries can be remediated through healing processes localized to the injury site. Furthermore, unlike the semantic frameworks for impurity, which indicate movement of a substance, the framework for harm involves one entity causing a change of state in another entity (Levin, 1993). Harm events have an intentional agent or stimulus who is assigned the property “cause”; the experiencer of the harm or affected patient is assigned a property indicating their resulting change of state (e.g., “subjugated,” “destroyed”). The dyadic morality model in which moral transgressions entail agents harming patients (e.g., Gray et al., 2012), fits well with the semantic framework of harm; it does not fit with the semantic frameworks for impurity, for which the agent-patient dyad is irrelevant.

In Study 1 (a-c), we designed experiments which capitalized on this prior work in verb semantics. We examined harm and impurity by analyzing words used to directly convey these concepts, and we tested whether these words in their active and passive forms are applied differently to agents and patients, respectively. One form of the words, passive participles (adjectival forms of verbs that end in *-ed*), are typically considered “patient-oriented,” and applicable to previously affected entities. The second, active participles (end in *-ing*), are considered “agent-oriented,” and applicable to entities having progressive effects (Haspelmath, 1994). To investigate how people understand humans as agents and patients of contamination

DISSOCIATING IMPURITY AND HARM

and injury, we asked participants to assign active participles (“contaminating or tainting”, “injuring or wounding”), and passive participles (“contaminated or tainted”, “injured or wounded”) to generic “victims” and “perpetrators” of crimes. This design let us assess how role in transgression (victim vs. perpetrator) and descriptor type (active/passive) affect the assignment of harm and impurity terms.

We expected (1) congruent assignment of active and passive participles for contamination, indicating its perceived active, progressive nature even in a passive state – for example, the *OED* definition for “tainted” specifies being “affected with some corrupting influence”. In the case of injury, (2) we did not expect congruent assignment of active and passive participles, showing dissociation between these concepts: whereas contamination implies that being contaminated involves being affected by a negative substance that transfers to others in a full, irreversible way (i.e., contaminating others), being injured does not. We also expected (3) that perpetrators and victims would be assigned the descriptors in different ways: perpetrators might qualify as contaminated, contaminating, and injuring *but* not injured; victims would be most likely to be considered injured, followed by contaminated. Finally, prior work had also investigated assignment of the passive participles for contamination (“contaminated or tainted”) and injury (“injured or wounded”) to victims; the more victims were viewed as contaminated, the less they were viewed as injured (Niemi & Young, 2016; Niemi et al., in press). Following on this prior research, we also tested the relationship between contamination and injury ratings with correlational analyses.

Study 1a

DISSOCIATING IMPURITY AND HARM

In Study 1a, we examined our hypotheses that (1) participants would assign active and passive participles for contamination equally, (2) participants would assign active participles for injury to perpetrators and passive participles to victims, and (3) that in participants' ratings perpetrators would qualify as "contaminated," "contaminating," "injuring" *but not* "injured" in participants' ratings, whereas victims would be most likely to be considered "injured", followed by "contaminated." Finally, we tested the relationships between assignment of injury and contamination participles.

Study 1a: Method

Participants were 153 individuals on Amazon's Mechanical Turk; 27 participants were excluded for not completing the study or not providing their worker identification number. We aimed for approximately 50 individuals per condition (active/passive participle) based on standards for reasonable sample size, given that we did not have similar studies to refer to as precedents. Participants took the study for a small payment; our final sample after these exclusions was 126 ($Mage(SD)=34.63(12.13)$; 62 female, 63 male, 1 selected other).

Participants were asked about both hypothetical perpetrators and victims of crimes. Descriptor type (active participle "-ing", passive participle "-ed") was varied between subjects. Prompts were in the form: "Please consider the following hypothetical crime {victim; perpetrator}: a {victim; perpetrator} of {crime}." The crimes were molestation, rape, strangling, and stabbing.³ In the passive participle condition, participants (n=63) were asked: "How injured or wounded is this person?" and "How contaminated or tainted is this person?" in

³ We were also interested in whether crime *type* [i.e., sexual (rape, molestation) versus non-sexual (stabbing, strangling)] would influence contamination ratings. We found small effects suggestive of this which are not our focus here. See Supplementary Material for additional analyses broken down by crime type.

DISSOCIATING IMPURITY AND HARM

counterbalanced order. In the active participle condition, participants ($n=63$) were asked: “How injuring or wounding is this person?” and “How contaminating or tainting is this person?” in counterbalanced order. Participants provided their ratings using labeled sliding scales with the instructions: “Please use the slider to indicate your response, [0 = Not at all] to [7 = Very much]”. Data and materials for this and all studies are available at <https://github.com/BLINDED/>. We report all measures, manipulations, and exclusions in these studies.

Study 1a: Results

We conducted an analysis of variance on contamination and injury ratings of perpetrators and victims in the active and passive descriptor conditions. We observed main effects for rating type ($F(1,124) = 4.12, p=.04, \eta^2 = .032$) and role ($F(1,124) = 30.78, p<.001, \eta^2 = .199$). We observed two-way interactions of descriptor condition and role ($F(1,124) = 109.32, p<.001, \eta^2 = .47$); descriptor condition and rating ($F(1,124) = 12.84, p<.001, \eta^2 = .09$); and role and rating ($F(1,124) = 96.49, p<.001, \eta^2 = .438$). The three-way interaction of rating, role, descriptor condition interaction was significant ($F(1,124) = 40.24, p<.001, \eta^2 = .24$). Participants rated perpetrators highly and equivalently “contaminating”, “contaminated”, and “injuring”, whereas they rated victims highly “injured,” in particular (see Figure 1 and descriptive statistics in Table 1).

Correlations among ratings. In prior work involving a richer experimental context that included more detailed crime vignettes, participants rated victims as “contaminated” and “injured” – these ratings were negatively correlated (Niemi & Young, 2016). Here, in the passive participle condition, “injured” and “contaminated” ratings were unrelated for victims ($r = -.106, p=.41$) and perpetrators ($r = .05, p=.69$). However, in the active participle condition, “injuring”

DISSOCIATING IMPURITY AND HARM

and “contaminating” ratings were positively correlated for victims ($r = .588, p < .001$) and perpetrators ($r = .308, p < .014$), suggesting that being actively contaminating is consistent with being actively injurious.

Summary. The findings of Study 1a (Figure 1 and Table 1) support the notion that people perceive contamination, unlike injury, to fundamentally involve active transfer of negative effects. (1) Consistent with our first hypothesis, “contaminating” was assigned equivalently with “contaminated.” (2) In line with our second hypothesis, the active and passive participles for injury were assigned differently: “injuring” and “injured” traded off. (3) Finally, as expected, perpetrators and victims were assigned the descriptors in different ways: perpetrators were “contaminated,” “contaminating,” and “injuring” *but not* “injured”; victims were considered “injured”. In sum, being contaminated is considered to be consistent with being contaminating and injuring, but not being passively injured. In order to ascertain whether people would be influenced in their assignment of the active participles “contaminating” and “injuring” if they were *also* asked for the passive participle “contaminated” and “injured” ratings, we administered the next study (Study 1b) with the descriptor condition varied within-subjects.

DISSOCIATING IMPURITY AND HARM

Table 1. Descriptive Statistics for Study 1 (a-c)

Role	Descriptor	Mean	SE	95% Confidence Interval	
				<i>Lower Bound</i>	<i>Upper Bound</i>
Perpetrator	Contaminating	5.74	0.23	5.28	6.20
		4.92	0.21	4.50	5.35
		5.73	0.28	5.18	6.29
	Contaminated	5.14	0.23	4.68	5.61
		4.92	0.20	4.53	5.32
		5.14	0.28	4.58	5.70
	Injuring	5.85	0.26	5.34	6.35
		5.03	0.24	4.55	5.51
		6.34	0.28	5.80	6.89
	Injured	1.97	0.26	1.46	2.48
		2.75	0.26	2.24	3.26
		2.79	0.28	2.24	3.34
Victim	Contaminating	1.83	0.30	1.24	2.41
		2.11	0.23	1.65	2.57
		2.51	0.28	1.96	3.05
	Contaminated	3.18	0.30	2.60	3.76
		2.58	0.23	2.11	3.04
		2.75	0.25	2.25	3.25
	Injuring	3.21	0.26	2.69	3.73
		3.61	0.28	3.06	4.17
		4.70	0.27	4.16	5.23
	Injured	5.94	0.26	5.42	6.46
		5.80	0.16	5.49	6.12
		5.89	0.25	5.40	6.37

Note. Statistics for each condition in the following order Study 1a, 1b, 1c.

DISSOCIATING IMPURITY AND HARM

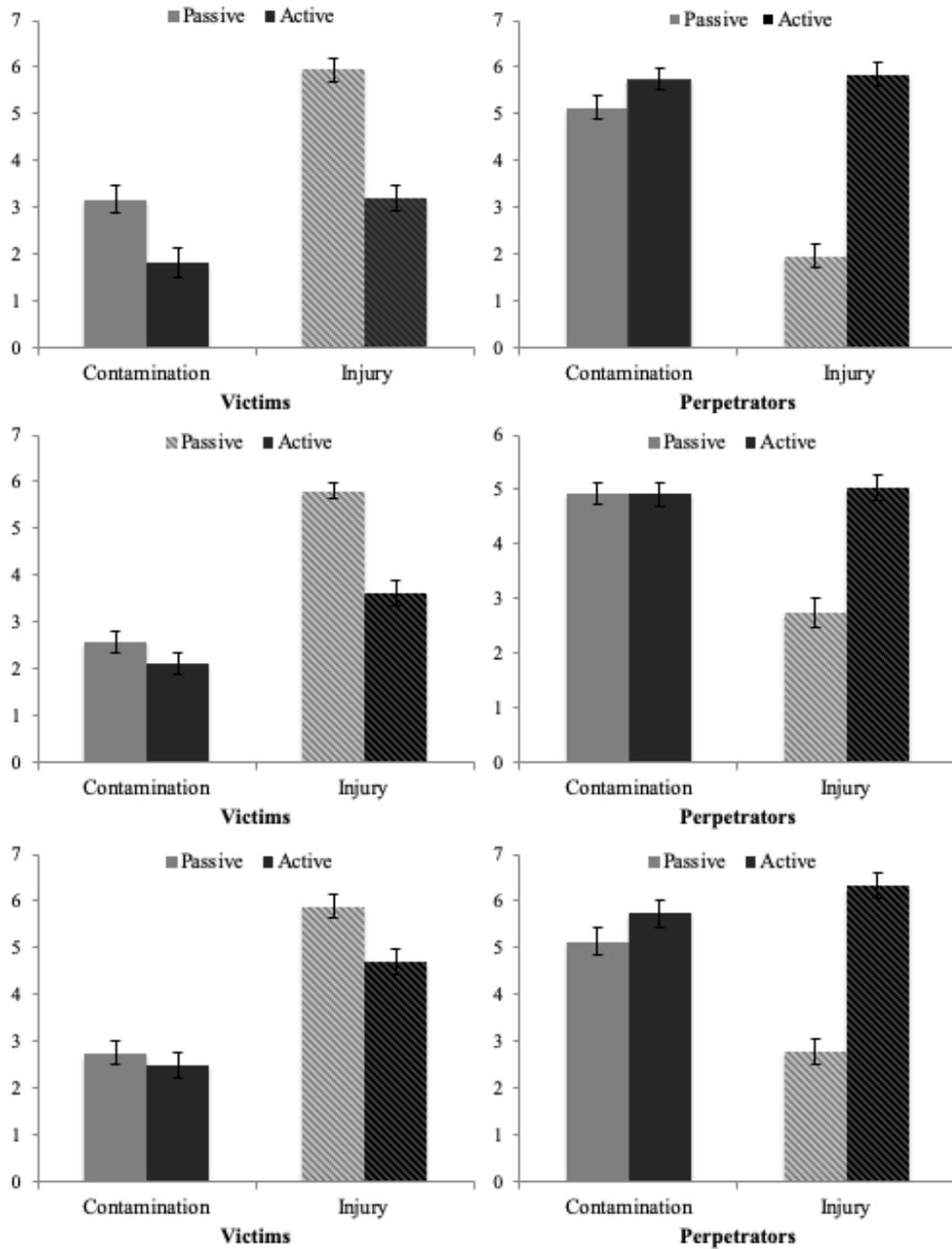


Figure 1. Ratings of contamination and injury (“not at all” - “very much”) in the active participle (“contaminating” and “injuring”) and passive participle (“contaminated” and “injured”) conditions for victims and perpetrators. Pattern of results was maintained across Study 1a (TOP) with active or passive participle condition varied between-subjects; Study 1b (MIDDLE) with active or passive particle condition within-subjects; and Study 1c (BOTTOM) with active or passive participle condition and victim / perpetrator role conditions between subjects. Error bars indicate SEM.

Study 1b

Study 1b: Method

In Study 1b, we asked: would asking participants to assign *both* the active and passive forms of the descriptors to victims and perpetrators eliminate the congruence in assignment of “contaminating” and “contaminated”? Study 1b was identical to Study 1a with the exception that conditions were varied within-subjects: participants completed both descriptor conditions (active/passive) for victims and perpetrators in randomized order. Participants included 154 individuals on Amazon’s Mechanical Turk who took the study for a small payment; 49 participants were excluded for not completing the study or for not providing their worker identification number. Our final sample comprised 105 individuals ($Mage(SD)=36.94(12.01)$; 55 female, 50 male).

Study 1b: Results

We conducted an analysis of variance on contamination and injury ratings of perpetrators and victims in the active and passive descriptor conditions. We observed main effects for rating type ($F(1,104) = 19.78, p < .001, \eta^2 = .160$) and role ($F(1,104) = 14.62, p < .001, \eta^2 = .123$); and two-way interactions of descriptor condition and role ($F(1,104) = 62.18, p < .001, \eta^2 = .37$) and role and rating ($F(1,104) = 106.78, p < .001, \eta^2 = .507$). The three-way interaction of rating, role, descriptor condition interaction was significant ($F(1,104) = 65.72, p < .001, \eta^2 = .387$). Participants again rated perpetrators highly and equivalently “contaminating”, “contaminated”, and “injuring”, whereas they rated victims highly “injured” (see Figure 1 and descriptive statistics in Table 1).

DISSOCIATING IMPURITY AND HARM

Correlations among ratings. Ratings of victims as “contaminated” were correlated with “contaminating” ($r = .72, p < .001$) and “injuring” ratings ($r = .22, p = .03$), but not “injured” ratings ($p = .12$). Likewise, ratings of perpetrators as “contaminated” were correlated with “contaminating” ($r = .71, p < .001$) and “injuring” ratings ($r = .48, p < .001$), but not “injured” ratings ($p = .06$).

Summary. In Study 1b, results replicated Study 1a. Giving participants both active and passive descriptors (e.g., “contaminated” and “contaminating”) in the same study did not change how they applied these terms. Again, supporting our first and second hypotheses, (1) participants assigned active and passive participles for contamination equally, and (2) participants assigned active participles for injury to perpetrators and passive participles for injury to victims. Also, (3) victims were again mostly considered “injured”, while perpetrators were considered “contaminated”, “contaminating” and “injuring” to a substantially great degree than victims. Correlations were found among “contaminated”, “contaminating”, and “injuring” ratings – but these were not correlated with “injured” ratings, for both victims and perpetrators.

Overall, these results provide further support for the notion that being passively contaminated is considered more likely to co-occur with being actively contaminating and injuring, but not passively injured. In our final iteration, Study 1c, we varied between-subjects whether participants rated victims *or* perpetrators, to rule out the possibility that the relationships between the descriptors may have been affected by comparisons made between “perpetrators” or “victims” when they were in the same study.

Study 1c

Study 1c: Method

DISSOCIATING IMPURITY AND HARM

Study 1c was identical to Study 1a except participants rated *either* victims or perpetrators rather than both. We doubled the sample size to accommodate the fully between-subjects design.

Participants were 198 individuals on Amazon's Mechanical Turk who completed the study for a small payment ($Mage(SD)=35.39(11.77)$; 102 female, 94 male, 2 other).

Study 1c: Results

We conducted an analysis of variance on contamination and injury ratings of perpetrators and victims in the active and passive descriptor conditions. We observed main effects for rating type ($F(1,194) = 27.17, p < .001, \eta^2 = .123$); descriptor condition ($F(1,194) = 10.46, p = .001, \eta^2 = .051$); and role ($F(1,194) = 24.71, p < .001, \eta^2 = .113$). We again observed two-way interactions of descriptor condition and role ($F(1,194) = 44.32, p < .001, \eta^2 = .19$); role and rating ($F(1,194) = 105.60, p < .001, \eta^2 = .352$); and descriptor condition and rating ($F(1,194) = 8.57, p = .004, \eta^2 = .04$). The three-way interaction of rating, role, descriptor condition interaction was significant ($F(1,194) = 32.19, p < .001, \eta^2 = .142$; see Figure 1 and descriptive statistics in Table 1): participants rated perpetrators more contaminating *and* contaminated than victims; whereas they rated perpetrators as more injuring than victims, and victims as more injured than perpetrators.⁴

Correlations among ratings. As in Study 1a, in the passive participle conditions for victims and perpetrators, “injured” and “contaminated” ratings were unrelated ($r = .025, p = .86$; $r = -.014, p = .93$). In the active participle condition, “injuring” and “contaminating” ratings were again positively correlated for victims ($r = .476, p < .001$), but not for perpetrators ($r = .04$,

⁴ Unexpectedly, we see moderately higher values for “injuring” for victims in Study 1c relative to Study 1a and Study 1b. Unlike the previous studies, in Study 1c, descriptor condition (“-ing” vs. “-ed”) was varied between-subjects, so participants saw either “injuring” and “injured”. Participants who saw “injuring” may have read it as they expected it would read for victims, as “injured”.

DISSOCIATING IMPURITY AND HARM

$p < .78$), partially replicating Study 1a and providing support for the notion that being actively contaminating is associated with being actively injurious.

Summary. The pattern of results in Study 1c, like Study 1(a-b), indicates (1) that people perceive passive contamination (being “contaminated”) to be consistent with active contamination (being “contaminating”); (2) there is no such active-passive congruency in the case of injury; and, (3) these descriptors are applied differently for victims versus perpetrators: victims were considered highly “injured” and substantially lower in the other descriptors, while perpetrators showed the opposite pattern.

The results support the conclusion that people understand impurity to fundamentally involve being not only passively contaminated but also actively contaminating: transferring the damage incurred — a feature not found in their understanding of harm. Moreover, across Study 1(a-c), participants assigned contamination terms to perpetrators more than victims, suggesting that perceptions of impurity are generally exacerbated by being an agent of harm.

Study 2 (a-b): Mechanics and Moral Implications of Contamination Judgments

Study 1 (a-c) provided linguistic evidence that the nature of impurity fundamentally differs from harm: being passively contaminated involves being actively contaminating to others. The results are consistent with research on contagion beliefs where tainted entities are perceived to have an essence that transfers upon contact to a target which then also becomes completely and irreversibly tainted (Rozin et al., 1986; Nemeroff & Rozin, 1994). The results are also consistent with a linguistic feature of impurity terms – they accommodate the grammatical construction that implies that a transferred substance filled its target, but not the grammatical construction that implies that it only affected its target in part (Levin, 1993). Thus, the evidence

DISSOCIATING IMPURITY AND HARM

so far is consistent with the theory, and strongly suggests that impurity diverges from harm because it entails inferences that an entity is completely and irreversibly affected and actively spreads negative effects to others. In Study 2 (a-b), we investigated how judgments of contamination and injury are affected by such inferences about the scope and duration of negative effects on a target (i.e., completeness, irreversibility, and contagious transfer). This time, we investigated the “perpetrator-victim” dyad by communicating these roles *indirectly*. The protagonist in a vignette was presented as either harmed (the victim) or harmed someone (the perpetrator). We manipulated completeness, irreversibility, and contagious transfer by varying the qualities of changes that the protagonist endorsed, which fit into these categories.

We also tested effects on blame judgments to investigate the moral implications of dissociating these domains. Prior work indicated that the passive participles for contamination and injury were differently related to responsibility judgments for victims: the more contaminated people judged victims, the more responsible they judged them; whereas the more injured they judged victims, the less responsible they judged them (Niemi & Young, 2016). By manipulating the features of changes the protagonist expressed and determining effects on contamination, injury, and blame ratings, Study 2 (a-b) aimed to both specify mechanisms of contamination judgments and elucidate their moral implications.

Study 2a

In Study 2a, we investigated how judgments of harm and impurity are affected by inferences about completeness, irreversibility, and transmission, prompted by a protagonist’s endorsements of personal changes representing these features. We used vignettes that conveyed the protagonists as a crime victim or perpetrator indirectly, as described next. We tested effects

DISSOCIATING IMPURITY AND HARM

on ratings of contamination, injury and blame, as well as on distractor variables including resilience and disrespect.

Study 2a: Method

Participants were 2400 individuals on Amazon's Mechanical Turk who completed the study for a small payment ($Mage(SD)=36.92(12.17)$; 1303 female, 814 male, 7 selected other, 2 n/a). 275 participants were excluded for not completing the study. Each participant read one of six vignettes featuring a quality hypothesized to relate to contamination (i.e., complete change $n=356$, partial change $n=355$; irreversible change $n=353$, reversible change $n=353$; transfer $n=353$, no transfer $n=355$). We were interested in whether each of these tested qualities would affect ratings of contamination and harm for a protagonist, Kim, presented either as a victim or as a perpetrator. In previous vignette research on blame, contamination, and injury, we found small relationships between ratings and moral values with samples of approximately three hundred participants. In the present study, we aimed for approximately three hundred participants per vignette to be able to detect potentially small effects with our subtler and briefer stimuli.

Status of the protagonist as a victim or perpetrator was implied. In the vignette where the protagonist was presented as a victim, "Kim was mugged" as in the vignette sample below, the term "victim" was not used. In the vignette where the protagonist was presented as a perpetrator, "Kim mugged someone," the term "perpetrator" was not used. The wording of each of the vignette variants (complete, partial, irreversible, reversible, transfer, no transfer) is provided in Table 2. Participants answered five questions about Kim after the vignette, presented in Table 3. The "**victim: complete change**" vignette was as follows:

DISSOCIATING IMPURITY AND HARM

Kim was mugged on April 7, 2013. After the court proceedings, Kim received a lot of treatment. As might be expected, all of this strongly affected Kim in many ways. Years later, Kim noted that “there was no part of herself that felt okay.”

Table 2. The altered portion of each vignette variants in Study 2a-b (all beginning “Years later, Kim noted that”).

Complete	... <i>“there was no part of herself that felt okay.”</i>
Partial	... <i>“some parts of herself felt okay, but other parts did not.”</i>
Irreversible	... she would <i>“never be the same again.”</i>
Reversible	... she <i>“was just starting to go back to how she used to be.”</i>
Transfer to others	... <i>“many people close to her were also affected.”</i>
No transfer to others	... <i>“no one close to her seemed to be affected.”</i>

Table 3. The items used in Study 2a-b.

- (1) How resilient is Kim, compared to other people you know?
- (2) To what extent was Kim harmed or injured?
- (3) To what extent was Kim tainted or contaminated?
- (4) To what extent was Kim disrespected or disobeyed?

How much blame do you believe Kim deserves for the mugging?

Note. Question 1 on resilience was presented first as a distractor item followed by questions 2-4 in randomized order; question 5 on blame was presented on a separate screen last. Responses were provided using a Likert scale anchored at 0 = “Not at all” and 7 = “Very much” for questions 1-4; and 0 = “None at all” and 7 = “A lot” for question 5. A short demographic survey completed the study.

Study 2a: Results

DISSOCIATING IMPURITY AND HARM

Kim as “tainted or contaminated”. We conducted an analysis of variance on ratings of contamination for the perpetrator and victim based on the tested qualities in the six vignette (complete, partial, irreversible, reversible, transfer, no transfer). We observed a main effect for role: victim Kim ($M(SEM) = 3.33(.07)$) was rated as slightly more “tainted or contaminated” than perpetrator Kim ($2.93(.07)$); $F(1,2075) = 17.06, p < .001, \eta^2 = .01$). We found a small main effect for the vignette qualities ($F(5,2075) = 3.85, p = .002, \eta^2 = .01$).

We observed a main effect of vignette type ($F(5,2075) = 3.85, p = .002, \eta^2 = .01$). Bonferroni-corrected follow-up comparisons indicated that people’s ratings of Kim as contaminated were increased relative to the other vignettes when she expressed that “there was no part of herself that felt okay”, i.e., complete change ($3.55(.12)$) compared to the other vignette types: transfer ($2.99(.12), p = .014$); no transfer ($2.96(.12), p = .006$); reversible ($2.95(.12), p = .005$); and partial variant (trend: $3.09(.12), p = .09$). Ratings of Kim as contaminated did not differ between the vignettes conveying complete and irreversible change. There was no significant difference in these effects based on her implied role as victim or perpetrator.

Kim as “harmed or injured”. We conducted the same analysis on ratings of Kim as “harmed or injured”. We observed a main effect of role ($F(1,2083) = 1790.23, p < .001, \eta^2 = .46$): consistent with Study 1(a-c), and unsurprisingly, victim Kim ($M(SEM) = 5.23(.05)$) was rated as substantially more “harmed or injured” than perpetrator Kim ($1.95(.069)$). There was no effect of the vignette type. Thus, the qualities hypothesized to be contamination-relevant had no effect on participants’ ratings of Kim as “harmed or injured”.

Kim as “disrespected or disobeyed”. We conducted the same analysis on ratings of Kim as “disrespected or disobeyed”. We observed a main effect of role ($F(1,2073) = 1337.53, p < .001, \eta^2 = .39$). Unsurprisingly, victim Kim ($5.24(.06)$) was rated as substantially more

DISSOCIATING IMPURITY AND HARM

“disrespected or disobeyed” than perpetrator Kim (1.92(.06)). There was a main effect of the vignette type ($F(5,2073) = 3.23, p < .007, \eta^2 = .008$) which Bonferroni-corrected pairwise comparisons indicated was driven by ratings of Kim as more disrespected in the partial variant relative to the no transfer variant ($p = .04$).

Kim as blameworthy. Regression analyses examined the contributions of ratings of Kim as harmed, disrespected, and contaminated on ratings of blame. For both Kim the perpetrator and Kim the victim, increased contamination ratings ($\beta = .166, p < .001$; $\beta = .136, p < .001$), reduced injured ratings ($\beta = -.356, p < .001$; $\beta = -.203, p < .001$), and reduced disrespected ratings ($\beta = -.283, p < .001$; $\beta = -.156, p < .001$) significantly contributed to increased blame ratings ($F(3,994) = 132.42, p < .001, R^2 = .29$; $F(3,1048) = 31.60, p < .001, R^2 = .08$, respectively). Viewing Kim as more tainted (and less injured and disrespected), whether she was the victim or perpetrator, contributed to an impression of her as more blameworthy.

Study 2b

Study 2b: Method

In Study 2a, results suggested that inferences about complete and irreversible change might affect judgments of a person as contaminated, but not harmed. In Study 2b, we tested the role of the same qualities (i.e., complete, partial, irreversible, reversible, transfer, no transfer) for “Kim” presented as either the victim ($n = 411$) or the perpetrator ($n = 406$) of a mugging event. This time, we included three of the qualities in each vignette, as opposed to including just one quality in each vignette in Study 2a. This presentation aligns with people’s typical, multi-dimensional self-presentations, which might describe changes in several of these qualities. This approach also let us test replication of the results of Study 2a, in a slightly modified design.

DISSOCIATING IMPURITY AND HARM

Participants were 817 individuals on Amazon's Mechanical Turk who completed the study for a small payment ($Mage(SD)=35.6(11.6)$; 484 female, 323 male, 2 selected other, 8 n/a). The sample size was reduced, as we changed the design from Study 2a where each vignette represented one of the six tested qualities for either a victim or perpetrator, to each vignette including one of each of the pairs of qualities (complete / partial, irreversible / reversible, transfer / no transfer) for a victim or perpetrator. For example, the vignette read by subjects for one of the variants: "victim: complete, irreversible, transfer" was as follows.

Kim was mugged on April 7, 2013. After the court proceedings, Kim received a lot of treatment. As might be expected, all of this strongly affected Kim in many ways. Years later, Kim noted that "there was no part of herself that felt okay," that she would "never be the same again," and that "many people close to her were also affected."

Following the vignette, participants were presented with same questions as in Study 2a.

Study 2b: Results

Kim as "tainted or contaminated". We conducted an analysis of variance on ratings of contamination for the perpetrator and victim based on the tested qualities (complete / partial, irreversible / reversible, transfer / no transfer). As in Study 2a, we observed a main effect for role: victim Kim ($3.71(.11)$) was again rated as slightly more "tainted or contaminated" than perpetrator Kim ($3.24(.11)$); $F(1,798) = 8.92, p=.003, \eta^2 = .01$). We observed an interaction of role with the qualities irreversible / reversible ($F(1,798) = 6.98, p=.008, \eta^2 = .01$). Victim Kim was rated as more contaminated when she expressed she would "never be the same again"

DISSOCIATING IMPURITY AND HARM

(irreversible; 4.00(.16)) compared to when she expressed she was “just starting to go back to how she used to be” (reversible; 3.42(.16)). This indicates that the victim’s irreversible change was likely perceived as being permanently negatively affected, causing judgments of contamination to increase. Perpetrator Kim was rated as *less* contaminated when she expressed that she would “never be the same again” (irreversible; 3.11(.16)) compared to when she was “just starting to go back to how she used to be” (reversible; 3.37(.16)). This finding suggests that the perpetrator’s irreversible change was likely perceived as a permanent positive change which caused judgments of contamination to decrease.

Kim as “harmed or injured”. We conducted the same analysis on ratings of Kim as “harmed or injured”. We observed a main effect of role ($F(1,798) = 452.10, p < .001, \eta^2 = .36$). Again, victim Kim (5.36(.09)) was rated as substantially more “harmed or injured” than perpetrator Kim (2.57(.09)). There was no effect of the tested qualities on ratings of Kim as “harmed or injured”.

Kim as “disrespected or disobeyed”. We conducted the same analysis on ratings of Kim as “disrespected or disobeyed”. We observed a main effect of role: victim Kim (5.50(.10)) was rated as substantially more “disrespected or disobeyed” than perpetrator Kim (2.25(.10); $F(1,798) = 528.50, p < .001, \eta^2 = .40$). There was no effect of the tested qualities on ratings of Kim as “disrespected or disobeyed”, indicating the small effect observed in Study 2a is not robust.

Kim as blameworthy. Regression analyses examined the contributions of ratings of Kim as harmed, disrespected, and contaminated on ratings of blame. Identical to Study 2a, for both Kim the perpetrator and Kim the victim, increased contamination ratings ($\beta = .187, p < .001$; $\beta = .145, p = .004$), reduced injured ratings ($\beta = -.341, p < .001$; $\beta = -.178, p < .001$), and reduced

DISSOCIATING IMPURITY AND HARM

disrespected ratings ($\beta = -.312, p < .001$; ($\beta = -.129, p < .001$) significantly contributed to increased blame ratings ($F(3,388) = 52.37, p < .001, R^2 = .29$; $F(3,403) = 8.85, p < .001, R^2 = .06$). In sum, viewing Kim as more tainted (and less injured and disrespected), whether she was the perpetrator or victim, contributed to an impression of her as blameworthy.

Summary. Across Study 2a-b, inferences posited to underlie contagion beliefs in linguistics and psychological science – namely, complete and irreversible change – affected judgments of a person as contaminated, but not injured. Contamination ratings predicted blameworthiness, regardless of the protagonist's role as harm-doer or harmed person.

General Discussion

In this research, we combined language analysis and the methods of social and moral psychology methods to investigate the features that distinguish impurity and harm, and their moral implications. We proposed that, while impurity and harm both involve negative consequences for the affected, impurity affects the target completely, leaves an irreversible mark, and progressively transfers to others. These features do not characterize harm. Across studies, we found evidence of the defining features of impurity – completeness, irreversibility, and transfer to others – in patterns of language use (Study 1 a-c) and in studies that measured the effects of manipulating these features on judgments of contamination and injury (Study 2 a-b).

First, harm and impurity are delineated by very different linguistic features. As discussed in our analysis of lexical semantics, linguists specify basic differences in meaning for verbs that convey harm and impurity. Harm terms accommodate a semantic framework in which a causal, intentional agent causes a state change in an affected patient, whereas impurity terms accommodate semantic frameworks in which a substance occupies or changes its location

DISSOCIATING IMPURITY AND HARM

(Levin, 1993). The model of moral cognition that specifies transgressions as events in which agents harm patients (e.g., Gray et al., 2012) fits well with the semantic framework of harm terms; it does not fit with the semantic frameworks for impurity terms for which the transfer of substance, rather than the agent-patient dyad, are relevant. Furthermore, grammatically, the verbs contaminate and taint can be used in constructions that imply that the target was completely affected (“holistic”), but not in constructions that allow for the interpretation that the target was only partially affected (“partitive,” Levin, 1993). By contrast, verbs conveying harm and injury are not characterized by this restriction. This grammatical feature is consistent with the hypothesis that impurity and harm dissociate in that impurity judgments, unlike harm judgments, communicate that an entity is negatively affected in entirety.

In Study 1 (a-c), we found, and twice replicated, differing patterns in how people apply participles, the adjectival form of verbs, for impurity and harm. People apply the active (“contaminating”) and passive (“contaminated”) participles for contamination nearly equivalently to targets, while they apply the active (“injuring”) and passive (“injured”) participles for injury differently, with “injuring” applied to perpetrators, and “injured” to victims. These findings support our hypothesis that a feature of impurity is an inference about progressive transfer: *to be contaminated is to be contaminating* – that is, patients of contamination by default are also agents of contamination. By contrast, in the case of injury, the active and passive participles cleanly map onto the agent / patient roles of perpetrator / victim. The linguistic results not only afford evidence that harm and impurity are distinct moral domains but also bring new specificity to our understanding of how and why harm and impurity psychologically diverge. While both are conditions involving damage to a target, they differ in that impurity implies that the affected individual is completely and irreversibly affected, in a way that can spread to others.

DISSOCIATING IMPURITY AND HARM

In our studies of language use, we also observed that people generally considered perpetrators higher in contamination compared to victims. This indicates that harmfulness plays a role in judgments of impurity. However, the contrasting behavior of the contamination and injury descriptors and the inferences they point to indicate that impurity cannot be conceptualized as merely another kind of harm. When impurity and harm seem to overlap it is likely due to a focus on their shared capacity to cause damaging effects. The qualities of those damaging effects, however, are crucial to the meaning of the concepts and are what separate the domains. The inference that an impure entity is completely damaged, is damaged for an indefinite duration, and becomes damaging to others starkly contrasts with harm, in that injuries and wounds are inferred to be isolable and to have the capacity to heal over time.

We further investigated the inferences driving impurity judgments in Study 2 (a-b). We precisely matched descriptions of a protagonist, “Kim”, presented as either a mugger or a mugged person, manipulated Kim’s everyday language statements about herself as experiencing complete, irreversible, and spreading negative effects from the event, and measured judgments of her as contaminated and harmed. The results were consistent with accounts from linguistics and psychological science that stipulate contagion cognition as involving beliefs about total and irreversible damage that spreads on contact (e.g., Rozin et al., 1986; Nemeroff & Rozin, 1994). Summarizing across Studies 2a and 2b, we found that when Kim was presented as completely and irreversibly changed, this affected participants’ contamination – but not harm – ratings, relative to changes including that her experiences affected others (i.e., transfer). Although contamination is a process evoking contagiousness, the protagonist’s statements about personal change of a complete and irreversible nature had effects on perceptions of contamination that statements implying spreading social effects did not. Future work should explore alternative

DISSOCIATING IMPURITY AND HARM

approaches to measuring these concepts using everyday language to better understand whether statements about spreading effects should be understood to be less crucial to social-moral inferences about contamination than statements implying irreversible and complete change.

Contamination of Harmed People and Harm-Doers

In prior exploratory work, the core concepts of complete and irreversible affectedness also appeared when people were asked to write freely about what they meant when they rated victims as “*contaminated or tainted*” (see a sample of responses in Supplementary Materials). Their responses describe a long-lasting temporal dimension to victim contamination, and complete, total change, e.g., in the case of rape: “*nothing will be the same*”, “*every aspect of the rape victim’s life is negatively affected*”, the victim “*may never be able to get over it*”. When people described contaminated victims, they described personal, often psychological, features they believe to be affected; less often did they specify how a person would affect other people by being contaminated. This suggests that evaluations of people as contaminated might be more likely to be based on how long in duration and how complete the negative effects appear to be for an affected individual, rather than whether they transfer negative effects to others. Alternatively, thinking about victims as contaminated may be construed by participants as a form of victim-blaming. Previous research indicates that judging victims as contaminated is consistent with victim-blaming. In this work, the more participants judged victims as contaminated, the less they considered victims injured, even though one might think that evaluations of all kinds of negative impacts to victims should track together (Niemi & Young, 2016; Niemi et al., in press). Furthermore, the more participants judged victims as contaminated, the more they considered victims to be causal contributors and deserving of harm (Niemi et al., in press). Aversion to

DISSOCIATING IMPURITY AND HARM

victim-blaming might reduce the salience of contamination when evaluating victims – in particular, the active aspect of contamination, as studied here, where victims spread damage to others.

Past research shows that some people are more likely than others to consider victims to be contaminated. People high in binding values, a cluster of group-oriented moral values (e.g., Graham et al., 2011), are more likely to judge victims as contaminated, causal, responsible, and blameworthy (Niemi & Young, 2016; Niemi et al., in press). As noted earlier, both the semantic structure of contamination and many of the transgressions under the purview of binding values do not fit with the dyadic framework. To understand the nature of contamination judgments in the context of individual differences in moral values, future work should explore whether the extent to which one endorses binding values determines which aspects of contamination one references when judging victims as impure (i.e., complete, irreversible, or spreading damage).

It is notable that in Study 1 (a-c), perpetrators were rated more contaminated than victims, whereas in Study 2 (a-b) this pattern flipped: a victim was actually rated as slightly more contaminated than a perpetrator. The results of Study 1 (a-c) seem to suggest that perceived harmfulness drives perceived impurity, consistent with our finding of a correlation between “injuring” and “contaminating” ratings. However, we also found that impure entities are both passively contaminated (completely and irreversibly negatively affected) and actively contaminating (transfers these negative effects to others). As active and passive contamination are intertwined, victims may sometimes be considered actively contaminating, and perpetrators passively contaminated. Indeed, in Study 2 (a-b), the vignette emphasized patient-like qualities (Kim receives treatment, feels strong emotions), and Kim the perpetrator’s contamination ratings

DISSOCIATING IMPURITY AND HARM

were greatly reduced, even slightly below Kim the victim's ratings. This further indicates that contamination is modulated by factors other than harm.

Future research investigating whether indicators of agent and patient role modulate contamination may be useful in order to construct a more detailed explanation for moral judgments of impurity. For example, we might learn more about why and how contamination judgments are applied to both perpetrators and seemingly unlikely targets like victims. One possible model might sort the features of contamination into two dimensions: (i) *passive* – the extent to which the target is completely and irreversibly negatively affected, and (ii) *active* – the extent to which the target transfers negative effects to others. The active and passive dimensions of impurity might come into focus differently depending on context, for example, when judging agents such as perpetrators and patients such as victims. However, the agent-patient framework does not enable the full conceptualization of impurity, which entails being both passively contaminated and actively contaminating. Another approach to explaining patterns of contamination judgments might explore whether people are referring to two different kinds of contamination when they refer to perpetrators and victims as contaminated. For example, they might infer that perpetrators and victims are contaminated by different symbolic “substances” such as evil or filth (Elliott & Radomsky, 2012; Nemeroff & Rozin, 1994).

Finally, regardless of the target's harm-doer or harmed status, the more participants judged Kim to be “tainted or contaminated”, the more she was rated as blameworthy. This reveals that judgments of a person as “contaminated or tainted” are likely to be found alongside accusatory judgments, which may have problematic normative implications for people who have been harmed, in particular. Namely, victim-blaming may be facilitated when targets are viewed as tainted.

Linking Approaches to Study Morality

This research demonstrates how morality can serve as a hub to unite theory and methods from across disciplines, in this case linguistics and experimental social and moral psychology. As such, our results have significance for our understanding of language as well as social and moral psychology. Our experimental designs rely on theory about lexical semantics and grammatical structure, as well as social psychology theory on contagion beliefs. As a result of this merger, the findings broadly contribute to our understanding of the relevance of theories of verb semantics to social and moral judgment. For example, as discussed, effects on objects on the spatial dimension – ‘complete affectedness’ -- allow the verbs *contaminate* and *taint* to be classified along with the *fill* verbs (based on alternation behavior; Levin, 1993). It has been proposed that the notions of *complete* and *continuous* affectedness may be hard to disentangle for theories of verb semantics because of the confoundedness of space and time (Croft, 2012). In line with this, across Study 2 (a-b), manipulation of the temporal (irreversibility) and spatial dimensions affected ratings of Kim as contaminated. Thus, the current research links the study of language and moral psychology by demonstrating that core concepts that intermingle in theories of verb semantics – complete spatial affectedness and unbounded temporal affectedness – also intermingle in their contributions to people’s *morally* relevant judgments. Furthermore, the findings show that the moral domains impurity and harm may not be collapsed together if we pay mind to the cognitive-linguistic underpinnings of the processes of contamination and injury at their roots. Discovering the distinctions between these moral domains would not have been possible without morality serving as the center point that accommodates and fruitfully unites multiple approaches.

Conclusions

Whether and how impurity and harm should be distinguished has been disputed; some commentators view impurity as harm, while others contend they represent distinct normative domains. These questions are illuminated, we have shown, by combining the resources of distinct, complementary methodologies. Through a combination of language analysis and vignette studies, both the partial overlap and clear differences between impurity and harm judgments are revealed. The concepts partially overlap in that they both entail negative, damaging effects. However, impurity cannot be simply reduced to harm; impurity involves unique inferences that the damage is complete, irreversible, and transfers to others. These inferences are encoded in the words and grammatical constructions associated with contamination, not injury, and when conveyed in everyday speech in vignettes, they altered people's contamination, and not harm, ratings, which in turn were associated with exacerbated blame. These results are important for our understanding of lexical semantics and moral cognition, and have novel practical implications. They indicate that when a person communicates that *every* aspect of their life has been negatively affected or that they will *never* be the same again (i.e., complete and irreversible change), this may not only reflect damage to personal well-being (e.g., Ehlers & Clark, 2000); it may also be consequential to others' morally-relevant judgments of them as contaminated or pure.

DISSOCIATING IMPURITY AND HARM

ACKNOWLEDGMENTS: Alek Chakroff, John Doris, Joshua Hartshorne, Joshua Knobe, Steven Pinker, Josh Rottman, Patricia Tueme, European Association for Social Psychology attendees, Boston-Area Moral Cognition Group, SPSP2017 Language in Values Symposium attendees, SESP2016 attendees, and funding from the John Templeton Foundation.

References

- Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., ... & Scelza, B. A. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, *113*(17), 4688-4693.
- Chakroff, A., Dungan, J., & Young, L. (2013). Harming ourselves and defiling others: what determines a moral domain?. *PloS ONE*, *8*(9), e74434.
- 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. & Young, L. (2015). Harmful situations, impure people: An attribution asymmetry across moral domains. *Cognition*, *136*, 30-37.
- Clifford, S., & Jerit, J. (2013). How words do the work of politics: Moral foundations theory and the debate over stem cell research. *The Journal of Politics*, *75*(3), 659-671.
- Croft, 2012. *Verbs: Aspect and Causal Structure*. Oxford University Press: Oxford, UK.
- Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., ... & Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, *145*(3), 366.
- Dungan, J. A., 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).
- Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, *38*(4), 319-345.

DISSOCIATING IMPURITY AND HARM

- Elliott, C. M., Radomsky, A. S. (2012). Mental contamination: The effects of imagined physical dirt and immoral behaviour. *Behavior Research and Therapy*, 50(6):422-7.
- 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, 366-385.
- Graham, J., Haidt, J., Motyl, M., Meindl, P., Iskiwitch, C., & Mooijman, M. (2018). Moral Foundations Theory: On the advantages of moral pluralism over moral monism. In K. Gray & J. Graham (Eds.), *The Atlas of Moral Psychology*. New York: Guilford
- Gray, K., Keeney, J. E. (2015). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological & Personality Science*, 6, 859–868.
- Gray, K., Schein, C., & Ward, A. (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.
- Haidt, J. (2003). The moral emotions. In R. Davidson, J. K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852-870). Oxford: Oxford University Press.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998-1002. doi: 10.1126/science.1137651
- Hildreth, J. A. D., Gino, F., & Bazerman, M. (2016). Blind loyalty? When group loyalty makes us see evil or engage in it. *Organizational Behavior and Human Decision Processes*, 132, 16-36.
- Kipper, K., Korhonen, A., Ryant, N. et al. (2008). A large-scale classification of English verbs.

DISSOCIATING IMPURITY AND HARM

- Language Resources & Evaluation*, 42, 21–40.
- Kipper-Schuler, K. (2005). *VerbNet: a broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, Computer and Information Science Department, University of Pennsylvania, PA.
- Haspelmath, M. (1994). Passive participles across languages. In: B. Fox & P. J. Hopper (Eds.) *Voice: Form and Function*. John Benjamins Publishing Company: Amsterdam/Philadelphia.
- Levin, B. (1993). *English Verb Classes and Alternations*. Chicago: Chicago University Press.
- Mooijman, M., Meindl, P., Oyserman, D., Monterosso, J., Dehghani, M., Doris, J. M., & Graham, J. (2018). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology*, 115(3), 585–599. Napier & Luguri, 2013
- Napier, J. L., & Luguri, J. B. (2013). Moral mind-sets: Abstract thinking increases a preference for “individualizing” over “binding” moral foundations. *Social Psychological and Personality Science*, 4(6), 754–759.
- Nemeroff, C., & Rozin, P. (1994). The contagion concept in adult thinking in the United States. *Ethos*, 22(2), 158-186.
- Niemi, L., Hartshorne, J., Gerstenberg, T., Stanley, M. & Young, L. (in press). Moral values reveal the causality implicit in verb meaning. *Cognitive Science*.
- Niemi, L., & Young, L. (2013). Caring across boundaries versus keeping boundaries intact: Links between moral values and interpersonal orientations. *PloS One*, 8(12).
- Niemi, L. & Young, L. (2016). When and why we see victims as responsible: The impact of ideology on attitudes toward victims. *Personality and Social Psychology Bulletin*, 42(9),

DISSOCIATING IMPURITY AND HARM

1227 – 1242.

- Niemi, L., Wasserman, E., & Young, L. (2018). The behavioral and neural signatures of distinct conceptions of fairness. *Social Neuroscience, 13*(4), 399-415.
- Parkinson, M., & Byrne, R. M. (2017). Judgments of moral responsibility and wrongness for intentional and accidental harm and purity violations. *The Quarterly Journal of Experimental Psychology, 1-12*.
- Rottman, J., & Young, L. (2019). Specks of dirt and tons of pain: Dosage distinguishes impurity from harm. *Psychological Science, 30*(8), 1151-1160.
- Rozin, P., Millman, L., & Nemeroff, C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology, 50*(4), 703-712.
- Russell, P.S., Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*.
- Sagi, E., & Dehghani, M. (2014). Measuring moral rhetoric in text. *Social Science Computer Review, 32*(2), 132-144.
- Schein, C., Ritter, R. S., & Gray, K. (2016). Harm mediates the disgust-immorality link. *Emotion, 16*(6), 862.
- 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.
- Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness–loyalty tradeoff. *Journal of Experimental Social Psychology, 49*(6), 1027-1033.
- Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition, 120*(2), 202-214. doi: 10.1016/j.cognition.2011.04.005