Article

Which Appraisals Are Foundational to Moral Judgment? Harm, Injustice, and Beyond

Jared Piazza[1] $\circledcirc$, Paulo Sousa[2], Joshua Rottman[3],
and Stylianos Syropoulos[3]

Abstract

Harm-centric accounts of judgments of moral wrongdoing argue that moral judgments are fundamentally based on appraisals of harm. However, past research has failed to operationally discriminate harm appraisals from appraisals related to injustice. Four studies carefully discriminated harm qua pain/suffering from injustice, alongside appraisals related to impurity, authority, and disloyalty. Appraisals of injustice outperformed appraisals of harm as independent predictors of the judged wrongness of recalled offenses (Study 1). Studies 2a, 2b, and 3 extended these findings using a diverse range of wrongful acts and two different cultural samples—the United States and Greece. In addition to the strong relevance of injustice appraisals, these latter studies uncovered substantial contributions of impurity and authority appraisals. The results inform debates on moral pluralism and the foundations of moral cognition.

Keywords

Perceptions of harm are important for moral judgments. But is perceived pain or suffering the fundamental input driving our judgments of moral wrongdoing? "Harm-centric" approaches to moral cognition posit that when people judge any act to be morally wrong, it is because they perceive the act to cause harm. In this regard, harm constitutes a foundational, organizing template by which all immoral actions are conceptualized (Gray & Schein, 2012; Gray, Schein, & Ward, 2014; Gray, Young, & Waytz, 2012; Schein & Gray, 2015, 2018).

By contrast, some have defended a deflationary view of harm, claiming that perceptions of harm cannot be sufficient for judgments of wrongdoing because people often find harmful acts acceptable (Fiske & Rai, 2014; Piazza & Sousa, 2016; Piazza, Sousa, & Holbrook, 2013; Sousa, Holbrook, & Piazza, 2009; Sousa & Piazza, 2014). When malevolent criminals are jailed, when a country attacks another country in self-defense, when professional boxers pummel each other in an arena, and when scientists subject animals to painful medical tests to test a vaccine, individuals are made to suffer. Yet, for many of us, these represent instances of acceptable harmful acts. Thus, appraisals beyond the causation of pain/suffering must be shaping our judgments of wrongdoing. This argument obtains even when harm is defined more broadly as welfare reduction, which is not necessarily tied to pain/suffering as a psychological state, or when restricting the definition of harm to the *intentional* causation of pain/suffering.

One increasingly popular harm-centric perspective is that of Gray and colleagues. This group of researchers sometimes characterizes harm simply in terms of the causation of pain/suffering (e.g., "judgments of harm require seeing a mind capable of suffering"; Schein & Gray, 2015, p. 3), yet other times they define harm more specifically as the intentional causation of pain/suffering: "harm involves the perception of two interacting minds, one mind (an agent) intentionally causing suffering to another mind (a patient)" (Schein & Gray, 2015, p. 3). Arguably, however, intentionality is not sufficient to elevate the causation of pain/suffering to the level of wrongdoing, as the example of punishment makes clear: People often think of punishment as deserved and therefore not wrongdoing.

The deflationary perspective on harm posits that if a harmful act is appraised as involving injustice, then it is judged to be morally wrong (for a detailed discussion, see Sousa & Piazza,

[1] Department of Psychology, Fylde College, Lancaster University, Lancaster, Lancashire, United Kingdom
[2] Institute of Cognition and Culture, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom
[3] Department of Psychology, Franklin & Marshall College, Lancaster, PA, USA

Corresponding Author:
Jared Piazza, Department of Psychology, Fylde College, Lancaster University, Lancaster, Lancashire LA1 4YF, United Kingdom.
Email: j.piazza@lancaster.ac.uk

2014; also Baumard, 2016). In this regard, the appraisal that a harmful act involves injustice *is* the appraisal that the actor did not consider the balance of interests involved when causing pain/suffering. Such an appraisal *prototypically* entails a belief that the actor acted from selfish motives: The actor either prioritized his or her own interests over that of others (e.g., when one steals from another person) or he or she preferentially prioritized the interests of another when fair treatment is expected (e.g., when a father gives preferential treatment to one of his children over another simply because he likes one more). Thus, in this regard, appraisals of injustice are generally linked to appraisals of selfishness.

Although deflationary theories claim that harm perception is insufficient for moral judgments of wrongdoing, they are neutral on whether perceptions of harm and/or injustice are necessary for such judgments. By contrast, pluralistic approaches to moral judgment, such as Shweder, Much, Mahapatra, and Park's (1997) "big three" ethical codes and Graham et al.'s (2013) moral foundations theory (MFT), argue that harm and injustice are not the only inputs to our judgments of wrongdoing. For example, Haidt (2007, 2012) has claimed that socially "binding" concerns related to respect for authority, loyalty to one's in-group (family, country, etc.), and the purity or sanctity of the body constitute distinct foundational sources of moral judgment.

Here, we report studies that shed new light on these debates concerning which appraisals are foundational to moral judgments of wrongdoing by addressing some methodological issues and limitations with previous research. First, we attempted to improve upon past operational definitions of harm. Schein and Gray (2015), for example, reported seven studies purportedly showing the foundational role of harm in moral judgment by using terms like "harm," "harmful," or related synonyms (e.g., "cruel") to operationalize the relevant concept of harm. However, these ordinary terms for harm are polysemous and often imply wrongdoing or even injustice. Indeed, this conflation of harm and wrongdoing can be observed in Schein and Gray's Study 1, which scored prototypically unjust acts, such as murder, stealing, and adultery, as forms of "harm." It remains unclear whether their results show that the relevant notion of harm, related to the causation of pain/suffering, is playing the key role in participants' judgments. Second, although deflationary theorists have made detailed theoretical arguments for their position (e.g., Sousa & Piazza, 2014), the evidence they have provided is mostly based on the reanalysis of other researchers' data (see, e.g., Piazza & Sousa, 2016). Moreover, they have not directly probed the role of injustice, which includes the perception of selfishness, in judgments of wrongdoing, nor have they systematically assessed its role across a diversity of transgressions beyond harmful transgressions.

Finally, although research by Graham, Haidt, and Nosek (2009) has arguably shown that many people do find concerns relating to impurity, disloyalty, and disrespect for authority to be relevant to their moral considerations, the validity of pluralistic approaches rests on demonstrating that each appraisal dimension contributes uniquely to judgments of wrongdoing and that each dimension is not perceived to be reducible to any other, for example, that the notion of impurity is not reducible to harm, as some have argued (Gray & Keeney, 2015). The present research design uniquely allowed us to measure the relevance of multiple appraisals for moral judgments across a wide range of moral transgressions that are representative of different moral foundations. We were therefore able to investigate the extent to which particular appraisals are consistently relevant for moral judgments across a diverse range of content versus having a restricted relevance. This approach provided a novel way of addressing prominent debates on moral monism and pluralism.

Overview of Studies and Hypotheses

We hypothesized that when care is taken to tease apart appraisals of causing pain/suffering and appraisals of injustice, the latter would provide a more extensive foundation for moral judgments. We also hypothesized, based on pluralistic theories, that other appraisal dimensions beyond harm and injustice (e.g., impurity) would make independent contributions to judgments of wrongdoing. In Study 1, American participants recalled an autobiographical experience of wrongdoing, rated its wrongfulness, and made 10 appraisals of the action. To allow an even broader test of moral pluralism, in the remaining studies, American (Studies 2a and 2b) and Greek participants (Study 3) were presented 10 transgressions related to Graham et al.'s (2013) five moral foundations. As in Study 1, participants judged their wrongness and appraised them. All collected measures and conditions are reported. Full materials and anonymized data sets for Studies 1–3 are available at https://osf.io/g7dpn/.

Study 1

In Study 1, we employed a recall paradigm that drew upon naturalistic perceptions of wrongdoing. The methodology was a revised version of that employed by Schein and Gray (2015) in their first study. Participants were asked to report a real instance of wrongdoing from their lives and rate its wrongness rather than abstractly "list an act that is morally wrong." Finally, for each reported act, participants were asked whether a series of 10 appraisals would apply to the act, and the applicability of each was measured on Likert-type scales (Schein and Gray's participants had to choose one of the five appraisals: "harmful, unfair, disloyal, disobedient, and gross"). Appraisals of harm were measured in terms of causing pain/suffering and welfare reduction ("The action caused someone pain," "The action negatively affected the well-being of someone"), separate from injustice ("The action was unjust," "unfair," "selfish"). We also included appraisals of impurity, disrespect for authority, and group disloyalty to connect to MFT (Graham et al., 2013). Finally, we included one appraisal, "dishonest" (also related to injustice), that is important in the literature on moral character (e.g., Brambilla & Leach, 2014; Goodwin, Piazza, & Rozin, 2014).

Method

Participants

We aimed to collect 160 adult participants on Amazon Mechanical Turk, restricting participation to those located in the United States and those who passed a captcha question; 160 falls within the sample size range required to determine whether a correlation coefficient at $r = .20–.25$ differs from zero with Type I error rate α (two-tailed) $= .05$ and Type II error rate $\beta = .20$ (Hulley, Cummings, Browner, Grady, & Newman, 2013). One hundred sixty-one workers completed the study and were paid US$0.50; five failed to provide any transgression or wrote nonsense (e.g., "gf"). One person failed to answer the wrongness probe. These participants were removed leaving $N = 155$ (85 male, 70 female; $M_{age} = 35.74$ years, $SD = 11.65$; 85% White/Caucasian, 7% Asian, 6% Black/African American, 2% Hispanic/Latino).

Materials and Procedures

After providing consent, participants were instructed: "We would like you to think about an action that you recently witnessed or heard about where someone did something wrong. This could be a minor offense or something major." They were given a large text box to describe the action. They were prompted to spend some time writing, and, to structure their response, were asked: "What was the person's relationship to you? What did they do? What was wrong about it?" The mean writing time was 2 min and 18 s ($SD = 209.83$). Next, on a separate page, participants rated the wrongness of the action on a 1–7 scale ($1 = not\ at\ all\ wrong$ to $7 = extremely\ wrong$). Then, they appraised the action on 10 dimensions, "to what extent do the following descriptions apply to the action you wrote about?" ($1 = not\ at\ all$ to $7 = extremely$). The 10 dimensions were: "The action . . . was unjust, selfish, unfair, dishonest, impure, made me feel nauseous (grossed out), was disrespectful toward an authority, involved someone being disloyal to their group, negatively affected the well-being of someone, caused someone pain." Finally, in all studies, participants answered basic demographic questions, were debriefed, and paid.

Data Reduction and Analysis Plan

We adopted a conceptual–empirical data reduction strategy that aggregated the items unjust, selfish, unfair, and *dishonest* (Cronbach's $\alpha = .85$) into a single injustice index, aggregated the items pain and negatively affected well-being ($\alpha = .80$) into a harm index and aggregated the items *impure* and nauseous (grossed out; $\alpha = .68$) into an impurity index. The single items related to authority (*disrespectful*) and group loyalty (*disloyalty*) were each treated separately. The same appraisal indices were used in all four studies (see Online Supplemental Materials for index reliabilities and exploratory factor analysis results).

In all four studies, we ran a linear regression on wrongness judgments using the appraisal indices (harm, injustice, impurity, disrespectful, disloyalty) as simultaneous predictors. In Studies 2a–3, as a secondary analytical strategy, we also conducted a mixed linear analysis of the five-factor model for each study to control for variability in the repeated judgments of participants across 10 scenarios and to take all five appraisals into account in a single analysis. See Online Supplemental Materials for additional analyses, and discussion, with the item "unjust" removed from the injustice index (as per the request of a reviewer).

Results and Discussion

Reported Wrongdoing

We first coded participants' qualitative responses ($N = 155$) to understand the diversity of moral content and to determine whether some responses were unscorable. The first author coded the responses and the fourth author coded them independently using the categories developed by the first author (Cohen's $\kappa = .725$). This two-rater procedure led to two original categories being dropped or merged with the others. Transgression categories are presented in Online Supplemental Figure S1. There was quite a diversity of transgressions reported (14 categories total); for 12% of responses, the nature of the transgression was unclear or unscorable.

Appraisal Ratings

The mean scores, and standard errors, for our five appraisal dimensions can be seen in Online Supplemental Figure S2. Note that while mean appraisal ratings offer some insight into the perceived relevance of each appraisal within a scenario, these ratings cannot answer the question of which appraisals contributed to variance in the wrongness ratings.

Main Analysis

Details regarding the distribution properties (skew, kurtosis) of the wrongness ratings and each predictor within the regression model can be found in the Online Supplemental Materials (see Online Supplemental Table S5 for distribution properties for all studies). Of the 138 scorable offenses, the mean wrongness rating was 5.81 ($SD = 1.28$). The five-factor model explained a significant amount of variance in participants' wrongness ratings, $R^2 = .59$, $F(5, 133) = 38.46$, $p < .001$. Injustice appraisals contributed the most predictive value, $\beta = .48$, $t(133) = 6.43$, $p < .001$, 95% confidence interval (CI) $= [.303, .573]$, followed by harm appraisals, $\beta = .29$, $t(133) = 4.08$, $p < .001$, 95% CI $[.098, .282]$. None of the other appraisals contributed significantly to wrongness judgments, βs $< .13$, ps $> .11$ (95% CIs contained 0). Thus, when we operationalized the concept of harm carefully (with terms related to pain/suffering and reduced welfare), we found that injustice provided a much stronger foundation for immorality. This was shown using a transgression recall paradigm that produced a large diversity of moral content (see Online Supplemental Figure S1).

In Study 1, we found little evidence for moral pluralism. However, certain immoral acts related to impurity, disrespect for authority, and group disloyalty (e.g., incest, betraying one's country) may be uncommon and thus may rarely appear in people's recollections of wrongdoing as prompted in Study 1. Thus, to deliberately cover the five foundations of morality articulated within MFT, Studies 2a and 2b presented participants with scenarios designed to evaluate wrongness judgments across five moral foundations, as theorized by Graham et al. (2013), thus, providing a wider test of moral pluralism.

Studies 2a and 2b

Studies 2a and 2b differed mainly in one aspect: Study 2a asked how "wrong" was each action, while Study 2b asked how "morally wrong" was each action. Study 2b used "morally" to address measurement commensurability with Schein and Gray (2015; see, e.g., Study 1), while Study 2a is consistent with the MFT approach, which avoids using the term "morally" in assessing judgments of wrongdoing within the Moral Foundations Questionnaire (see Graham, Haidt, & Nosek, 2009).

Method

Participants

We recruited two new samples of MTurk workers based in the United States and analyzed data from all individuals who completed the study and passed the captcha question. Participants were paid US$1.00. We aimed to recruit a minimum of 200 participants in each study. In Study 2a, 231 individuals started the survey and 206 completed it (124 male, 82 female; M_{age} = 34.87 years, SD = 11.20; 78% White, 11% Asian, 7% Black, 4% Hispanic or other). In Study 2b, 272 individual started the survey and 251 completed it (141 male, 110 female; M_{age} = 35.52 years, SD = 12.28; 76% White, 10% Black, 7% Asian, 7% Hispanic or other). Study 2a ran from April 16 to 22, 2016; Study 2b ran from June 9 to 18, 2018.

Materials and Procedures

The procedure was nearly identical for both studies. Participants provided informed consent and then completed 10 transgression blocks (two actions for each foundation) presented in a randomized order. Each scenario described a unique, female actor who engaged in a transgressive action (see Table 1). The scenarios were derived from Graham et al. (2009) but were modified to clarify the motive of the actor. This allowed transgressive aspects of the actions to be made explicit and discouraged participants from thinking that the actors may have had good reasons for engaging in the acts (e.g., in the dog scenario, inferences of self-defense were preempted by specifying that the kicking was motivated by dislike). For each scenario block in Study 2a, participants judged whether it was wrong or not wrong for the actor to have engaged in the act. If they selected

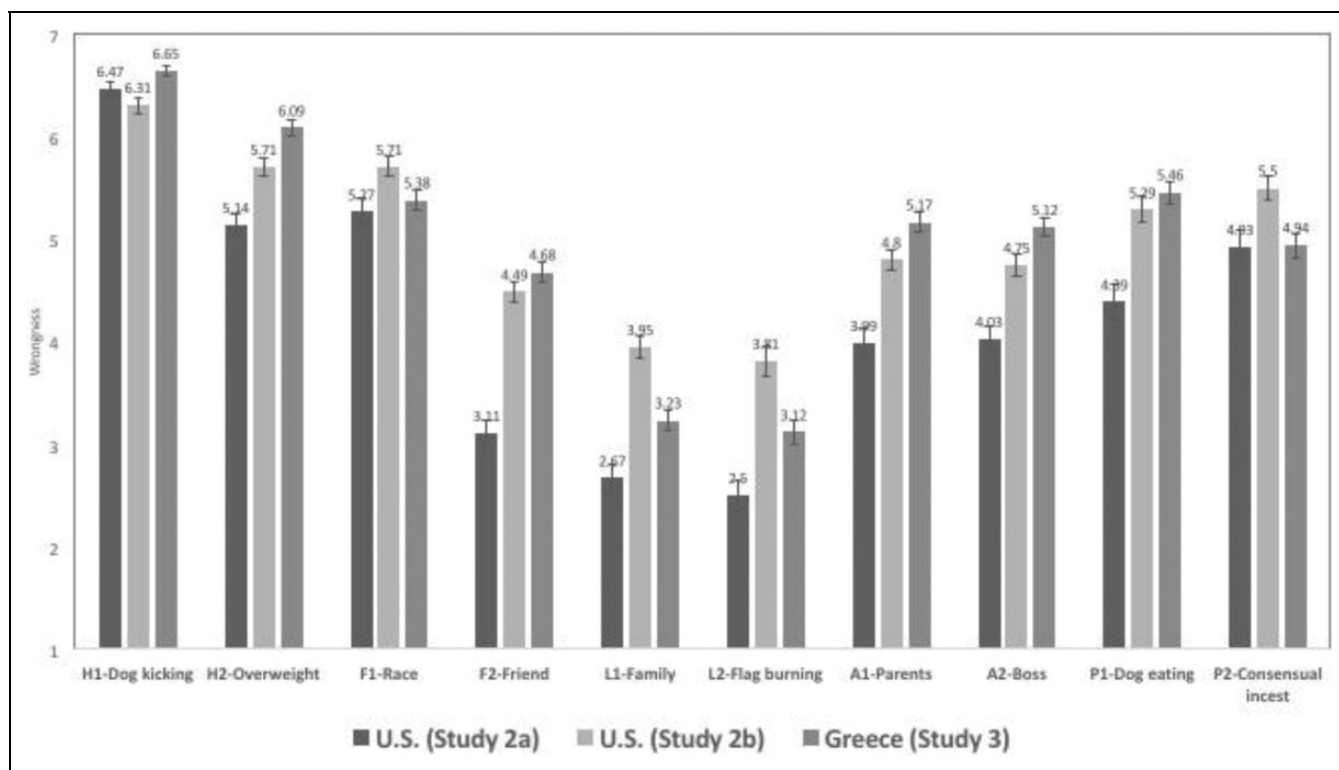**Table 1.** Transgression Scenarios Used in Studies 2a, 2b, and 3 .

Category	Scenario
Harm	(H1) Abby kicks a dog in the head, hard, because she doesn't like it.
	(H2) Cindy makes cruel remarks to an overweight person about their appearance because she doesn't like them.
Unfairness	(F1) Robin only hires people of her race in her company because she prefers working with people of her own race.
	(F2) Even though Maria has some free time, Maria does not help her friend move into a new apartment after her friend had helped her move the month before.
Group disloyalty	(L1) Nicola breaks off all communication with her immediate and extended family for one year simply because she had a heated argument with them.
	(L2) Fiona burns her country's flag in private (nobody else sees her) because she doesn't like her country.
Authority	(A1) Clare curses her parents to their face simply because she is angry with them.
	(A2) Jen makes a disrespectful hand gesture to her boss in a group meeting because she doesn't like her boss.
Impurity/ degradation	(P1) Lisa cooks and eats her dog after it dies of natural causes because she wanted to see what it tastes like.
	(P2) Julia has consensual sex with her biological brother who she has known all her life. They are both adults, desire each other, performed the act in private, and used contraceptives.

"wrong," they were then asked to rate how wrong (1 = *not wrong at all* to 7 = *extremely wrong*). The "*not wrong*" responses were scored 1. In Study 2b, the first step of this process was eliminated and participants simply rated how "morally wrong" the action was using the same 1–7 scale. Participants then rated the act on the 10 appraisal dimensions from Study 1 on the same 1–7 scale.

Results and Discussion

Wrongness and Appraisal Ratings

Figure 1 depicts the wrongness ratings for the 10 transgression scenarios for Studies 2a and 2b. There was a great deal of variation across the scenarios in the mean levels of judged wrongdoing (Study 2a: range = 2.50 [Loyalty 1] to 6.47 [Harm 1]; Study 2b: range = 3.81 [Loyalty 1] to 6.65 [Harm 1]). Although the means were slightly higher in Study 2b, perhaps due to the elimination of the initial binary probe, the pattern of means was consistent across studies. Mean ratings, and standard errors, of the five appraisal indices as a function of scenario can be found in Online Supplemental Figures S3 and S4.

Figure 1. Mean wrongness ratings and standard errors ($\pm$ 1 *SE*) for the 10 transgression scenarios used in Studies 2a and 2b (American sample) and Study 3 (Greek sample).

Main Analysis

Tables 2 and 3 show the results of the full regression model for the 10 scenarios, along with collinearity statistics (multicollinearity was not an issue except in one instance, flag burning, for both studies, predominantly for the injustice and impurity indices). In Study 2a, the injustice index was a significant predictor of wrongdoing for 9 out of 10 scenarios (β range = .15–.64), consensual incest being the exception. By contrast, the harm index was a significant predictor for only the two harm scenarios and one of the fairness scenarios (friend). The results for Study 2b were quite similar. The injustice index significantly predicted wrongdoing in all scenarios (β range = .29–.65). The harm index significantly predicted wrongdoing in 7 out of 10 scenarios, though it was a negative predictor in one of those seven (flag burning), and, with the exception of the harm scenarios, it was a weaker predictor than injustice. These findings show that injustice appraisals were quite foundational across moral diversity, whereas harm appraisals were much less foundational, though not insubstantial.

Consistent with moral pluralism, we observed domain-specific contributions from domain-relevant appraisals for nearly every moral foundation. In Study 2a, appraisals of group disloyalty contributed significantly to moral judgments of group disloyalty for both disloyalty scenarios (one in Study 2b) and appraisals of disrespect for authority contributed selectively to judgments of the authority scenarios (one in Study 2b). Appraisals of impurity were significant predictors for both

purity scenarios, but also had a wider contribution to other domains of action, including both harm scenarios, one of the fairness scenarios (friend), one of the disloyalty scenarios (flag burning), and one of the authority scenarios (parents); in Study 2b, impurity appraisals predicted wrongdoing judgments for six scenarios, including both purity scenarios.

Mixed Linear Model

To determine which of the appraisals impacted participants' wrongness judgments across the 10 scenarios, data for each study were analyzed with a linear mixed model fit with Satterthwaite approximation. The model was specified to predict wrongness judgments from the fixed effects, our five appraisal indices, and the random effects (intercepts) of participant and scenario. The results of the analysis converged with the results of the regressions. In Study 2a, the injustice index had the largest individual contribution overall, $B = .472$ ($SE = .031$), $t(1,939) = 15.01$, $p < .001$, 95% CI [.411, .533], followed closely by the purity index, $B = .394$ ($SE = .024$), $t(1,680) = 16.53$, $p < .001$, 95% CI [.347, .441], then the harm index, $B = .160$ ($SE = .026$), $t(1,742) = 6.22$, $p < .001$, 95% CI [.110, .210]. The contributions made by authority and disloyalty did not reach statistical significance, $B = .036$ ($SE = .020$), $t(1,822) = 1.75$, $p = .080$, 95% CI [−.003, .075], $B = .017$ ($SE = .021$), $t(1,962) = 0.83$, $p = .405$, 95% CI [−.024, .058], respectively. The results for Study 2b were quite similar to Study 2a: The injustice index had the largest individual

Table 2. Appraisals Predicting Judgments of Wrongdoing From Study 2a (American Sample Without Using "Morally").

Scenario	Appraisals	Model β	t	95% CI	R^2	Tolerance	VIF
Harm 1—Dog kicking					.60		
	Injustice index	**.15**	2.13*	[.009, .230]		.43	2.32
	Harm index	**.55**	10.56***	[.625, .913]		.75	1.34
	Impurity index	**.27**	4.34***	[.085, .226]		.54	1.86
	Disrespectful to authority	−.07	−0.99	[−.088, .029]		.47	2.14
	Disloyal to group	−.04	−0.51	[−.079, .046]		.40	2.51
Harm 2—Overweight					.50		
	Injustice index	**.33**	3.80***	[.184, .581]		.34	2.92
	Harm index	**.41**	6.37***	[.372, .706]		.61	1.64
	Impurity index	**.16**	2.29*	[.021, .278]		.49	2.03
	Disrespectful to authority	**−.17**	−2.34*	[−.269, −.023]		.51	1.97
	Disloyal to group	.01	0.17	[−.114, .136]		.46	2.19
Fairness 1—Race					.59		
	Injustice index	**.64**	9.32***	[.739, 1.135]		.44	2.29
	Harm index	.13	1.86	[−.010, .347]		.44	2.25
	Impurity index	.06	0.97	[−.064, .189]		.52	1.91
	Disrespectful to authority	−.04	−0.67	[−.156, .077]		.57	1.76
	Disloyal to group	.04	0.64	[−.072, .140]		.61	1.64
Fairness 2—Friend					.47		
	Injustice index	**.43**	5.71***	[.364, .747]		.46	2.16
	Harm index	**.17**	2.42*	[.034, .341]		.52	1.94
	Impurity index	**.21**	2.67**	[.068, .451]		.42	2.38
	Disrespectful to authority	−.10	−1.50	[−.303, .041]		.53	1.89
	Disloyal to group	.05	0.75	[−.080, .179]		.57	1.76
Loyalty 1—Family					.43		
	Injustice index	**.37**	3.77**	[.223, .751]		.31	3.28
	Harm index	.09	1.22	[−.075, .318]		.51	1.94
	Impurity index	.13	1.87	[−.011, .398]		.57	1.75
	Disrespectful to authority	.02	0.32	[−.122, .169]		.60	1.65
	Disloyal to group	**.17**	2.09*	[.011, .346]		.45	2.22
Loyalty 2—Flag burning					.60		
	Injustice index	**.36**	3.40**	[.180, .679]		.18	5.53
	Harm index	−.09	−1.25	[−.305, .069]		.37	2.71
	Impurity index	**.38**	4.12***	[.246, .696]		.24	4.24
	Disrespectful to authority	.00	0.00	[−.131, .131]		.45	2.24
	Disloyal to group	**.20**	2.82**	[.061, .342]		.39	2.53
Authority 1—Parents					.44		
	Injustice index	**.32**	3.52**	[.177, .627]		.34	2.91
	Harm index	.12	1.72	[−.024, .357]		.57	1.75
	Impurity index	**.18**	2.26*	[.027, .398]		.47	2.13
	Disrespectful to authority	**.25**	4.37***	[.197, .521]		.84	1.19
	Disloyal to group	.01	0.15	[−.147, .171]		.47	2.15
Authority 2—Boss					.40		
	Injustice index	**.32**	3.34**	[.149, .580]		.33	3.04
	Harm index	.12	1.59	[−.029, .275]		.56	1.77
	Impurity index	.11	1.37	[−.056, .307]		.48	2.10
	Disrespectful to authority	**.18**	3.15**	[.100, .435]		.90	1.11
	Disloyal to group	.12	1.51	[−.034, .261]		.51	1.96
Purity 1—Dog eating					.62		
	Injustice index	**.34**	4.32***	[.243, .652]		.30	3.35
	Harm index	−.11	−1.96	[−.375, .001]		.58	1.73
	Impurity index	**.58**	10.55***	[.665, .971]		.62	1.62
	Disrespectful to authority	.01	0.18	[−.180, .216]		.32	3.11
	Disloyal to group	−.02	−0.27	[−.221, .167]		.29	3.45
Purity 2—Consensual incest					.58		
	Injustice index	.06	0.62	[−.157, .300]		.26	3.78
	Harm index	.07	0.89	[−.098, .258]		.33	3.03
	Impurity index	**.66**	11.43***	[.719, 1.019]		.64	1.55
	Disrespectful to authority	.04	0.57	[−.109, .197]		.46	2.20
	Disloyal to group	.02	0.30	[−.121, .165]		.46	2.17

Note. Ns = 202–206. βs in boldface are significant at p < .05. R^2 given for full model. CI = confidence interval; VIF = Variance Inflation Factor.
*p < .05. **p < .01. ***p < .001.

Table 3. Appraisals Predicting Judgments of Wrongdoing From Study 2b (American Sample Using "Morally").

Scenario	Appraisals	Model			R^2	Tolerance	VIF
		β	t	95% CI			
Harm 1—Dog kicking					.56		
	Injustice index	**.41**	5.79***	[.267, .542]		.36	2.74
	Harm index	**.43**	7.93***	[.414, .687]		.62	1.61
	Impurity index	.07	1.16	[−.040, .156]		.44	2.25
	Disrespectful to authority	−.07	−1.12	[−.100, .027]		.47	2.12
	Disloyal to group	−.01	−0.19	[−.080, .066]		.38	2.60
Harm 2—Overweight					.56		
	Injustice index	**.43**	5.93***	[.266, .532]		.35	2.89
	Harm index	**.42**	7.83***	[.321, .537]		.63	1.59
	Impurity index	.11	1.63	[−.017, .178]		.37	2.70
	Disrespectful to authority	−.04	−0.68	[−.101, .049]		.52	1.92
	Disloyal to group	−.12	−1.74	[−.171, .010]		.36	2.82
Fairness 1—Race					.56		
	Injustice index	**.47**	6.90***	[.394, .708]		.40	2.53
	Harm index	**.26**	4.08***	[.140, .403]		.46	2.17
	Impurity index	.09	1.45	[−.026, .173]		.49	2.06
	Disrespectful to authority	−.01	−0.19	[−.085, .070]		.58	1.74
	Disloyal to group	.08	1.48	[−.019, .132]		.60	1.67
Fairness 2—Friend					.61		
	Injustice index	**.39**	5.93***	[.304, .606]		.37	2.69
	Harm index	**.19**	2.83**	[.057, .320]		.36	2.75
	Impurity index	**.30**	4.09***	[.135, .387]		.31	3.23
	Disrespectful to authority	.02	0.24	[−.088, .112]		.43	2.30
	Disloyal to group	.01	0.15	[−.092, .108]		.55	1.82
Loyalty 1—Family					.69		
	Injustice index	**.48**	6.34***	[.358, .681]		.23	4.39
	Harm index	**.11**	2.19*	[.013, .247]		.49	2.04
	Impurity index	**.31**	4.95***	[.171, .397]		.34	2.94
	Disrespectful to authority	−.06	−1.27	[−.138, .030]		.55	1.81
	Disloyal to group	.08	1.54	[−.022, .181]		.49	2.03
Loyalty 2—Flag burning					.79		
	Injustice index	**.65**	8.49***	[.561, .900]		.15	6.75
	Harm index	**−.18**	−2.89**	[−.313, −.059]		.22	4.56
	Impurity index	**.37**	5.27***	[.252, .553]		.18	5.61
	Disrespectful to authority	.01	0.14	[−.081, .094]		.50	1.99
	Disloyal to group	**.08**	2.01*	[.002, .170]		.51	1.95
Authority 1—Parents					.53		
	Injustice index	**.33**	3.89***	[.162, .494]		.27	3.66
	Harm index	.12	1.88	[−.006, .261]		.46	2.17
	Impurity index	**.28**	3.81***	[.119, .374]		.36	2.81
	Disrespectful to authority	**.15**	2.82**	[.046, .260]		.67	1.49
	Disloyal to group	−.02	−0.32	[−.126, .091]		.45	2.25
Authority 2—Boss					.57		
	Injustice index	**.54**	6.70***	[.392, .721]		.27	3.69
	Harm index	**.17**	2.72**	[.045, .281]		.46	2.18
	Impurity index	−.05	−0.71	[−.176, .083]		.34	2.96
	Disrespectful to authority	.08	1.72	[−.013, .196]		.73	1.36
	Disloyal to group	**.14**	2.43*	[.025, .234]		.54	1.85
Purity 1—Dog eating					.64		
	Injustice index	**.45**	5.75***	[.285, .581]		.25	4.08
	Harm index	.03	0.38	[−.097, .144]		.34	2.95
	Impurity index	**.49**	9.37***	[.462, .707]		.55	1.83
	Disrespectful to authority	**−.13**	−2.08*	[−.234, −.007]		.36	2.80
	Disloyal to group	.00	0.05	[−.111, .118]		.32	3.08
Purity 2—Consensual incest					.64		
	Injustice index	**.29**	3.48**	[.122, .441]		.21	4.83
	Harm index	−.01	−0.10	[−.131, .118]		.29	3.39
	Impurity index	**.59**	11.88***	[.530, .741]		.60	1.67
	Disrespectful to authority	−.09	−1.67	[−.160, .013]		.55	1.81
	Disloyal to group	.05	0.88	[−.056, .147]		.39	2.59

Note. Ns = 249–251. βs in boldface are significant at *p* < .05. R^2 given for full model. CI = confidence interval.
*p < .05. **p < .01. ***p < .001.

contribution overall, $B = .449$ ($SE = .025$), $t(2,457) = 17.98$, $p < .001$, 95% CI [.400, .498], followed closely by the purity index, $B = .319$ ($SE = .018$), $t(2,094) = 18.06$, $p < .001$, 95% CI [.284, .354], then the harm index, $B = .131$ ($SE = .019$), $t(2,234) = 7.00$, $p < .001$, 95% CI [.094, .168]. The contributions made by authority, $B = -.002$ ($SE = .014$), $t(2,295) = -0.16$, $p = .873$, 95% CI [−.029, .025], and disloyalty, $B = -.019$ ($SE = .015$), $t(2,441) = -1.29$, $p = .196$, 95% CI [−.048, .010], did not reach statistically significant levels.

Study 3

In Study 3, we sought to extend our findings to a different cultural context, Greece, which traditionally places great emphasis on familial bonds and parental discipline (Rosenthal, Bell, Demetriou, & Efklides, 1989), as an initial test of whether our claims about the extensive role of injustice, and moral pluralism, are culturally bounded.

Method

Participants

Our aim was to recruit at least 200 participants living in Greece. We obtained permission from the ethics committee at Aristotle University of Thessaloniki to circulate a web link to the study within a psychology classroom and a Facebook page that many students from the university frequent. This strategy led to a total 434 students who completed the entire survey (many others started but did not complete the survey). Among those who reached the end of the survey, 20 provided partial moral judgment responses at an unacceptable level (over three missing data points) or no demographic data. Thus, 414 participants were retained (113 males, 301 females; $M_{age} = 20.69$ years, $SD = 2.90$). Ninety-six percent of participants reported a Greek nationality, 99% White/Caucasian.

Materials and Procedures

The materials and procedures were identical to Study 2a. To obtain a Greek version, the fourth author first translated the English materials to Greek. This Greek version was then back translated by a second person proficient in Greek and English (see https://osf.io/g7dpn/).

Results and Discussion

Wrongness and Appraisal Ratings

Direct statistical comparisons of the two culturally distinct samples were deemed inappropriate and therefore not carried out. Relative to the U.S. sample from Study 2a, Greek participants had higher wrongness ratings for most of the transgressions, but this was less true compared to American Sample 2b, which at times had the highest wrongness ratings (see Figure 1). Like the U.S. sample, there was a great deal of variation between scenarios (means ranged from 3.12 [Loyalty 2] to 6.65

[Harm 1]). Mean appraisal ratings can be found in Online Supplemental Figure S5.

Main Analysis

We used the analysis strategy of Studies 2a and 2b. Table 4 shows the results of these analyses, along with collinearity statistics (there were no instances of multicollinearity). Quite similar to Studies 2a and 2b, the injustice index was a significant predictor for all 10 scenarios (β range = .27–.48), highlighting its foundational role. By contrast, the harm index was a significant predictor of wrongdoing for only 5 of the 10 scenarios—the two harm scenarios, one fairness (race), loyalty (family), and authority (parents; β range = .13–.25). We observed evidence of moral pluralism again. Disrespect-for-authority appraisals were a significant predictor of wrongness for 7 of the 10 scenarios, including both authority scenarios, and the impurity index significantly contributed to wrongness judgments for seven scenarios, including both purity scenarios. However, group-disloyalty appraisals had a small negative contribution to two scenarios (dog kicking, flag burning).

Mixed Linear Model

Similar to Studies 2a and 2b, injustice made the greatest contribution to wrongness judgments across domains, $B = .538$ ($SE = .024$), $t(4,022) = 22.08$, $p < .001$, 95% CI [.491, .585], followed by impurity, $B = .288$ ($SE = .019$), $t(3,568) = 15.10$, $p < .001$, 95% CI [.251, .325]. Yet differently, authority appraisals made a substantive independent contribution, $B = .147$ ($SE = .017$), $t(4,051) = 8.75$, $p < .001$, 95% CI [.114, .180], which was even greater than the contribution made by the harm index, $B = .103$ ($SE = .018$), $t(3,140) = 5.59$, $p < .001$, 95% CI [.068, .138]. Like Studies 2a and 2b, group disloyalty did not substantially contribute to wrongness judgments across content, $B = -.022$ ($SE = .014$), $t(4,044) = -1.55$, $p = .122$, 95% CI [−.049, .005].

Thus, once again, we observed evidence for a pluralistic account of moral judgment. While appraisals of injustice served as a ubiquitous foundation for wrongdoing in the Greek sample, there was substantial evidence as well for individual contributions made by other dimensions, including appraisals of harm, impurity, and disrespect for authority.

General Discussion

Across four studies incorporating two different methodologies and two different nationalities (American, Greek), we found that injustice appraisals provided a conceptual foundation for judgments of moral wrongdoing that was unmatched by harm appraisals. Furthermore, we found evidence for moral pluralism. Study 1 focused on offenses grounded in participants' real experiences, which produced a diverse range of content, with stealing, deception, killing, and rudeness as the most common transgressions (see Online Supplemental Figure S1). Despite this moral diversity, injustice and harm were the only

Table 4. Appraisals Predicting Judgments of Wrongdoing From Study 3 (Greek Sample).

		Model					
Scenario	Appraisals	β	t	95% CI	R^2	Tolerance	VIF
Harm 1—Dog kicking					.30		
	Injustice index	**.28**	4.76***	[.157, .378]		.50	1.98
	Harm index	**.21**	4.21***	[.153, .421]		.75	1.34
	Impurity index	**.11**	2.12*	[.005, .125]		.63	1.57
	Disrespectful to authority	**.17**	3.26**	[.039, .159]		.67	1.50
	Disloyal to group	**−.10**	−2.00*	[−.090, −.001]		.71	1.41
Harm 2—Overweight					.42		
	Injustice index	**.33**	6.01***	[.255, .503]		.50	1.98
	Harm index	**.25**	5.27***	[.221, .485]		.64	1.55
	Impurity index	.10	1.80	[−.007, .166]		.52	1.94
	Disrespectful to authority	**.16**	3.50**	[.055, .194]		.71	1.40
	Disloyal to group	−.05	−0.95	[−.112, .039]		.61	1.64
Fairness 1—Race					.47		
	Injustice index	**.40**	6.69***	[.441, .808]		.39	2.59
	Harm index	**.13**	2.57*	[.045, .339]		.53	1.89
	Impurity index	**.15**	2.95**	[.055, .275]		.52	1.91
	Disrespectful to authority	**.13**	2.48*	[.030, .261]		.53	1.90
	Disloyal to group	−.00	−0.10	[−.092, .083]		.69	1.45
Fairness 2—Friend					.41		
	Injustice index	**.44**	8.04***	[.483, .795]		.50	2.00
	Harm index	.09	1.78	[−.012, .257]		.58	1.72
	Impurity index	**.13**	2.54*	[.035, .274]		.59	1.69
	Disrespectful to authority	.09	1.89	[−.004, .178]		.67	1.49
	Disloyal to group	.00	0.05	[−.096, .101]		.67	1.49
Loyalty 1—Family					.39		
	Injustice index	**.39**	6.18***	[.458, .885]		.40	2.50
	Harm index	**.18**	3.78***	[.148, .470]		.68	1.47
	Impurity index	.08	1.57	[−.034, .304]		.55	1.82
	Disrespectful to authority	.08	1.43	[−.034, .214]		.52	1.91
	Disloyal to group	.01	0.14	[−.113, .130]		.61	1.64
Loyalty 2—Flag burning					.54		
	Injustice index	**.48**	8.07***	[.483, .794]		.33	3.04
	Harm index	.07	1.62	[−.020, .213]		.58	1.73
	Impurity index	**.21**	3.70***	[.131, .429]		.36	2.81
	Disrespectful to authority	**.13**	2.87**	[.051, .270]		.54	1.85
	Disloyal to group	**−.09**	−2.09*	[−.208, −.007]		.57	1.75
Authority 1—Parents					.40		
	Injustice index	**.34**	5.81***	[.304, .616]		.45	2.20
	Harm index	**.15**	3.12**	[.093, .406]		.65	1.53
	Impurity index	**.14**	2.62**	[.039, .270]		.53	1.88
	Disrespectful to authority	**.19**	3.80***	[.108, .340]		.64	1.55
	Disloyal to group	−.07	−1.36	[−.160, .029]		.66	1.52
Authority 2—Boss					.40		
	Injustice index	**.45**	7.78***	[.450, .755]		.46	2.17
	Harm index	.03	0.71	[−.071, .151]		.68	1.48
	Impurity index	.04	0.82	[−.075, .182]		.62	1.61
	Disrespectful to authority	**.26**	5.36***	[.213, .460]		.68	1.47
	Disloyal to group	−.07	−1.41	[−.165, .027]		.65	1.53
Purity 1—Dog eating					.48		
	Injustice index	**.33**	5.78***	[.268, .544]		.41	2.41
	Harm index	−.03	−0.61	[−.122, .064]		.80	1.23
	Impurity index	**.41**	8.69***	[.428, .678]		.58	1.71
	Disrespectful to authority	.06	1.27	[−.042, .193]		.50	2.01
	Disloyal to group	−.01	−0.17	[−.098, .082]		.62	1.62
Purity 2—Consensual incest					.55		
	Injustice index	**.27**	4.67***	[.211, .519]		.35	2.86
	Harm index	.07	1.54	[−.021, .179]		.56	1.79
	Impurity index	**.42**	8.85***	[.412, .648]		.51	1.96
	Disrespectful to authority	**.15**	3.22**	[.072, .297]		.52	1.94
	Disloyal to group	−.06	−1.33	[−.160, .031]		.57	1.76

Note. Ns = 400–408. βs in boldface are significant at $p < .05$. R^2 given for full model. CI = confidence interval.
*$p < .05$. **$p < .01$. ***$p < .001$.

dimensions that independently predicted ratings of wrongness. Studies 2a, 2b, and 3 adopted an experimenter-driven methodology to ensure a wider coverage of content and sampled from the United States (Studies 2a and 2b) and Greece (Study 3). For both cultural samples, we found that injustice appraisals contributed significant, independent variance to all, or all but one, of the transgressions, highlighting its extensive, foundational role. Appraisals of harm contributed substantially to several transgressions, for both samples, but this contribution was much narrower than that of injustice appraisals. Critically for debates regarding moral pluralism, we found that several other appraisals—those related to appraisals of impurity (all three studies) and disrespect toward authority (for the Greek sample)—contributed to judgments of wrongdoing across diverse content.

When taken together, our findings provide substantial support for the role of injustice as a comprehensive foundation for moral judgment, certainly stronger and more far-reaching than appraisals of harm. Furthermore, our findings support a pluralistic view of moral judgment. Not only did appraisals distinct from harm and injustice predict moral judgments in sensible ways, at least one appraisal dimension (impurity) made a more extensive, independent contribution to moral judgments than monistic theories would predict (e.g., Gray & Keeney, 2015). The present findings advance an understanding of which appraisals are central to moral judgments of wrongdoing and which contribute more narrowly. While past research on the deflationary theory of harm has indicated that appraisals of injustice are essential to viewing a harmful action as transgressive (Piazza & Sousa, 2016; Piazza et al., 2013; Sousa et al., 2009), no research to date has shown injustice appraisals to be important to *all sorts of* wrongdoing, beyond those involving harm. Our methodological approach also advances work on the topic of moral pluralism, not because we find strong evidence of different "domains" of moral evaluation but instead because we demonstrate that several distinctive appraisals are implicated in a range of moral judgments.

One important limitation is that the methodologies employed here were not ideal for conducting principal component analyses because all 10 appraisals were negatively valenced. While two-factor structures were found across the three studies, the two components were difficult to interpret due to multiple cross-loadings (see Online Supplemental Materials for details). Future studies should put effort into developing alternative methodologies that are better suited for testing the conceptual boundaries between morally relevant appraisals—for example, factor analytic strategies that use semantic similarity–dissimilarity ratings (e.g., rate the similarity of these statements: "causing someone pain," "being disloyal to a group," etc.). Second, we measured rather than manipulated appraisals, limiting the causal conclusions we can draw. Finally, we measured moral wrongdoing without probing additional criteria, such as authority independence or generalizability, that could more clearly differentiate normative evaluations related to moral versus conventional transgressions (see Turiel, 1983).

Why did we find evidence for moral pluralism when some recent findings have found otherwise? The answer may have to do with divergent operationalizations of harm. Many studies that have challenged the moral relevance of appraisals such as unfairness, impurity, and disloyalty by pitting these dimensions against harm (see, e.g., Gray et al., 2014; Schein & Gray, 2015) have operationalized harm with words like "harmful" and "cruel"—which express not only harm qua pain/suffering but also the *unjust* causation of pain/suffering, thereby favoring terminologically the relationship between harm and immorality. Thus, we operationalized harm by using expressions related to pain and well-being that more unambiguously expressed the intended concept. Yet, this still leaves open the empirical possibility that many transgressions are considered immoral partly because they are perceived as involving the (intentional) causation of suffering or welfare reduction. Indeed, in our studies, appraisals of pain/welfare reduction did contribute to wrongness judgments across a range of immoral content. Nevertheless, our findings do not support the monistic argument that harm appraisals are the foundation of all immorality, as in many cases, other appraisal dimensions surpassed the contribution made by harm appraisals and/or harm appraisals did not contribute much at all.

Conclusion

Four studies showcased the foundational role of injustice for morality, beyond the role played by harm, and simultaneously revealed support for moral pluralism. Not all harmful actions are considered morally wrong, and our findings indicate that harm is not the foundation of all moral wrongdoing.

ORCID iD

Jared Piazza ![ORCID] https://orcid.org/0000-0001-7261-3939

Supplemental Material

The supplemental material is available in the online version of the article.

References

Baumard, N. (2016). *The origins of fairness: How evolution explains our moral nature*. Oxford, England: Oxford University Press.

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition*, *32*, 397–408.

Fiske, A. P., & Rai, T. S. (2014). *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge, England: Cambridge University Press.

Goodwin, G., Piazza, J., & Rozin, P. (2014). Moral character information predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*, 148–168.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55–130.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*, 1029–1046.

Gray, K., & Keeney, J. E. (2015). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*, *6*, 859–868.

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*, 405–423.

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

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101–124.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002.

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

Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D., & Newman, T. B. (2013). *Designing clinical research: An epidemiologic approach* (4th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Piazza, J., & Sousa, P. (2016). When injustice is at stake, moral judgements are not parochial. Proceedings from the Royal Society of London B, *283*. doi:10.1098/rspb.2015.2037

Piazza, J., Sousa, P., & Holbrook, C. (2013). Authority dependence and judgments of utilitarian harm. *Cognition*, *128*, 261–270.

Rosenthal, D. A., Bell, R., Demetriou, A., & Efklides, A. (1989). From collectivism to individualism? The acculturation of Greek immigrants in Australia. *International Journal of Psychology*, *24*, 57–71.

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

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*, 32–70.

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

Sousa, P., Holbrook, C., & Piazza, J. (2009). The morality of harm. *Cognition*, *113*, 80–92.

Sousa, P., & Piazza, J. (2014). Harmful transgressions qua moral transgressions: A deflationary view. *Thinking & Reasoning*, *20*, 99–128.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, England: Cambridge University Press.

Author Biographies

Jared Piazza is a lecturer of psychology at Lancaster University. His research focuses on questions in moral psychology, such as the role of emotion and religion in the way people approach moral issues and factors influencing how we treat animals.

Paulo Sousa is a director of the Institute of Cognition and Culture at Queen's University Belfast. His current research interests focus on agency, moral psychology, and intergroup conflict as well as their relation to religion.

Joshua Rottman is an assistant professor of psychology and scientific and philosophical studies of mind at Franklin & Marshall College. His research lies at the intersection of social cognitive development and moral psychology, focusing on issues such as children's acquisition of moral norms, the role of disgust in forming intergroup biases, and the cognitive foundations of environmentalist beliefs.

Stylianos Syropoulos is an undergraduate student majoring in psychology at Franklin & Marshall College. He works in the Developing Moral Values Lab under the supervision of Joshua Rottman. His research focuses on attributions of moral concern and on measuring perceptions of safety.

Handling Editor: Jesse Graham