When Good People Break Bad: Moral Impression Violations in Everyday Life

Kate W. Guan and Steven J. Heine

Department of Psychology, University of British Columbia

Abstract

The present research investigated the emotional, interpersonal, and impression-updating consequences of witnessing events that violate the moral character impressions people hold of others. Across three studies, moral character-violations predicted broad disruptions to participants' sense of meaning, confidence judging moral character, and expectations of others' moral characters. Participants who were in real life closer to perpetrators, directly victimized, and higher in preferences for closure and behavioral stability reported more negative outcomes. Moreover, experimental manipulations showed that character-violations lead to worse outcomes than the comparable experience of encountering consistently immoral others. The authors discuss implications for research on moral perception and meaning, as well as on understanding responses to everyday revelations about people's characters.

Keywords: morality, moral character, meaning, impression revision, closeness

When Good People Break Bad: Moral Impressions in Everyday Life

In 2018, a judge sentenced Bill Cosby to prison following explosive revelations of his decades of criminal misconduct. His crimes horrified a public who had idolized Bill Cosby as "America's dad." How did this clash in character perceptions, between the publicly moral Bill Cosby and the immorality of his actual behavior, affect onlookers? Why are people bothered by such apparent moral inconsistencies?

The Need for Meaning

People possess a fundamental need for meaning (Heine et al., 2006), and are motivated to construct and maintain coherent mental models of their environments. When new information contradicts these models, people experience the aversive arousal of a physiological threat response (Tullett et al., 2011). Even innocuous meaning-threats like nonsensical word-pairings (Randles et al., 2011), or playing cards with reversed colors (Randles et al., 2018) motivate people to engage in compensatory action to restore their sense that the world is comprehensible and predictable (Janoff-Bulman, 1992).

Humans are particularly motivated to maintain meaning in their models of their social environments (Heine et al., 2006). As a cultural species, humans live in complex relational structures in which they must understand, predict, and act alongside other agentic actors (Dunbar, 1993; Tomasello, 1999). To decrease the vast, and threatening, social uncertainty about other agents' intentions and likely behavior, humans are motivated to make quick assumptions about others' traits, goals, and motivations (FeldmanHall & Shenhav, 2019). Quick social inferences (e.g., Skowronski et al., 2008) and implicit theories about personality (Dweck et al., 1995; Levy et al., 2006) provide comfort that people are predictable. Conversely, when new information contradicts these theories, people feel a threatening loss of prediction and meaning

(Plaks et al., 2005). Analogous to an unexpectedly black queen of hearts (Randles et al., 2018), revelations of an unexpectedly malevolent Bill Cosby threaten the meaningfulness of people's valued social assumptions.

The Importance of Moral Character

Beyond implicating the social domain, we theorize that contradictory character perceptions pose even stronger threats to meaning compared to typical social expectancy-violations, due to the primacy of morality in social cognition (Goodwin et al., 2014; Leach et al., 2007). People are more concerned with assessing others' morality, than their competence or sociability (Goodwin et al., 2014), due to morality's implications for whether those agents will help or harm, be potential allies or enemies (Abele & Wojciszke, 2007; Cottrell et al., 2007). Indeed, people are motivated to judge morality quickly; they inquire about a stranger's moral qualities before any other kind (De Bruin & Van Lange, 2000), and they base global dispositional judgments on one's moral character (Goodwin et al., 2014). Assessments of moral character feel like assessments of who someone really is (Christy et al., 2017; Hartley et al., 2016), their innate essence (Heiphetz, 2019), and true identity (Strohminger et al., 2017; Strohminger & Nichols, 2014). Finally, these perceptions help organize broader perceptions about the social world, determining who has moral standing (Piazza et al., 2014), which groups are worth helping (Brambilla et al., 2012), and who should be approached or avoided (Brambilla et al., 2013).

Given the predictive and epistemic value of perceptions of moral character, these perceptions may play an important role in people's sense of a meaningful world (e.g., Just-World Theory; Lerner, 1980). In addition, other accounts have theorized that perceptions that others are morally good, and the world a moral world, can also be potent sources of meaning (Janoff-

Bulman, 2012). Character-violations may therefore threaten people's sense of meaning in uniquely potent ways.

Moral Character-Violations Threaten Our Core Assumptions

We propose that character-violations are uniquely potent as they violate both core assumptions about an agent's identity, as well as broader assumptions about living in a moral world (Janoff-Bulman, 2012). That is, character-violations can prompt doubt about whether there are other unseen, misunderstood social threats around: Do other "good" people also conceal unseen immorality? This kind of meaning threat may result in a strong sense of aversive arousal (Harmon-Jones et al., 2009) and fear of danger in one's surroundings (e.g., Harris & Fiske, 2010). Though past research suggests character-violations make people feel betrayed (Koehler & Gershoff, 2003), or conflicted between punishing or protecting the individual (Weidman et al., 2020), we propose an additional response: Revelations about a single other's perceived immoral nature may undermine people's broader epistemic confidence in their ability to judge character.

The Present Research

Across three studies, we tested this theory that character-violations broadly disrupt perceivers' sense of meaning and specifically affect their character judgments of unrelated others. That is, does the degree to which a "good" person violates impressions predict how much people feel a broad sense of meaning violation, doubt in their moral character judgments, and cynicism when it comes to people's characters in general? We focus on good-to-bad character-violations, given people's general bias to focus more on negative events (Rozin & Royzman, 2001), and evidence that participants see good-to-bad changes as bigger shifts in identity than bad-to-good changes (e.g., De Freitas et al., 2017).

Studies 1 and 2 employed a distinct methodological approach to examine meaning disruptions surrounding character-violations in people's everyday lives. This retrospective paradigm, which asked participants to describe their own past moral character-violation experiences, contrasts with typical de-contextualized paradigms in moral psychology (Bloom, 2011). It also enabled us to capture rich descriptive details and to explore how antecedents and moderators predict self-reported outcomes. Study 3 then experimentally induced character-violations to establish causality and directionality of effects. Full items for all studies are reported in the SOM, and additional exploratory measures, data, and R code are reported at: https://osf.io/vf8mc/?view_only=6a9d1cf86b0d458e8a4a939ef53d4eeb.

Study 1

Study 1 explored the kinds of past good-to-bad character-violations people reported having, and the relationship between these experiences and perceivers' sense of meaning, confidence assessing others' morality, and beliefs about the good character of other people.

Method

Participants

We aimed to sample 315 participants, giving us 95% power to detect small correlations ($r=.20$). Because the study was designed for participants who could recall past character-violations, we planned to screen out participants who could not recall such a memory. We therefore screened 600 MTurk United States residents, and screened out 166 participants who could not recall a character-violation: 434 (72%) remained. Finally, 373 participants were retained after passing two attention checks (which asked them to select specific scale items) and an English comprehension check (testing participants' understanding of a short passage) (39% male, $M=36.1$ ($SD=11.8$) years old).

Procedure

Participants first reported details about a past moral character-violation, such as the target's age, gender, and number of years since the event. The study then proceeded with:

Part 1 (Before-Violation). This section asked participants to respond from their point of view prior to the character-violation. Participants indicated their before-violation relationship with the target ("stranger," "acquaintance," "co-worker," "friend," "romantic partner," "family member," "public figure," or "other"), and character judgment of the target (7 statements: e.g., "[This person] is a person of strong moral character", 1 ("strongly disagree") to 7 ("strongly agree")). They then reported their before-violation moral impressions of other people (1-item: "What was your overall impression of people in general at that time?", 1 ("extremely negative") to 9 ("extremely positive")) and confidence judging character (1-item: "How would you have rated your abilities as a good judge of character at that time?", 1 ("extremely poor") to 9 ("extremely good")).

Part 2 (During-Violation). Measures in this section referred to participants' memories of the character-violation itself. We operationalized the character-violation as the degree of inconsistency between participants' before-violation impressions of the target and the new information they learned about the target ("To what degree was this 'new information' inconsistent with your initial impression of [this person]?", 1 ("not at all inconsistent") to 7 ("extremely inconsistent")). Participants reported their sense of meaning violation, using four items from the Global Meaning Violation Scale (GMVS) Belief Violations subscale (Park et al., 2016), e.g., "How much did the occurrence of this experience violate your sense of the world being fair or just?", 1 ("not at all") to 5 ("very much"). Participants again indicated their moral

impressions of other people and confidence judging character during this time, using the same

two items as in Part 1.

Part 3 (After-Violation). This section asked participants to respond from their

perspective following the character-violation event, using the same items as in Part 1.

Participants then completed a demographic questionnaire.

Results and Discussion

All continuous predictors are standardized. Means and correlations among all relevant variables

are summarized in Table 1.

Table 1

Means, Standard Deviations, and Correlations of all Relevant Variables in Study 1

Measures	*M*	*SD*	1	2	3	4	5
Part 1 (Before-Violation) Measures							
1 Character Judgment of Target (α = .92)	5.6	1.0	-				
Part 2 (During-Violation) Measures							
2 Inconsistency	6.0	1.7	.46***	-			
3 Meaning Violation (α = .84)	2.9	1.2	.08	.13*	-		
4 Confidence Judging Character	3.8	1.7	-.08	-.08	-.19***	-	
5 Moral Impressions of Others	4.3	2.0	.01	-.06	-.25***	.43***	-
Part 3 (After-Violation) Measures							
6 Character Judgment of Target (α = .94)	2.6	1.4	.05	-.15***	-.19***	.23***	.35***

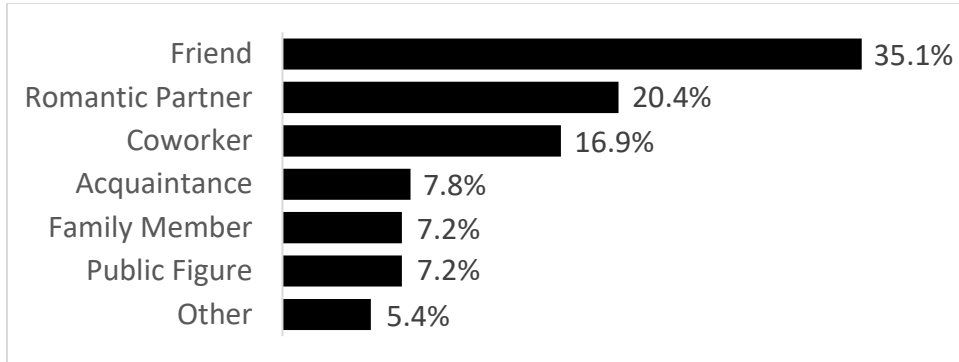*Note*. *p≤.05, **p≤.01, ***p≤.001.

Common Features of Character-Violations

Frequency and Descriptive Data. To start, we first explored common features of these

character-violations. The majority of participants screened (72%) recalled at least one distinct

character-violation experience. Figure 1 depicts the different types of targets described. Targets

were generally older than participants (*M*=3.5 years older, *SD*=13.0), and more commonly male

(62.5%) despite our sample being mostly female. These experiences modally occurred 1 year

after participants came to know targets (*M*=5.2 years, *SD*=7.4).

Figure 1

Frequency of Different Target Types in Study 1

        Impression-Change. As expected, perceivers' positive before-violation character

judgments of targets (*M*=5.6, *SD*=1.0) dropped sharply following character-violations (*M*=2.6,

SD=1.4). The degree of impression change, operationalized as after-violation character

judgments controlling for before-violation judgments, directly tracked the perceived

inconsistency of target's behaviors (β=-0.30, *SE*=0.08, 95% CI [-0.46, -0.15], *p*<.001).

Disruptions in Social Perception

        We analyzed whether moral character-violations predicted participants' sense of meaning

and impressions of others' moral characters, controlling for baseline Time 1 scores where

measured (see Table 2). Results showed that the more inconsistent targets' behaviors appeared,

the more participants experienced meaning violations (β=0.15, 95% CI [0.03, 0.26]). Moreover,

target inconsistency during these character-violations predicted participants feeling less

confident judging moral character during the time these events took place, controlling for Part 1

confidence (β=-0.19, 95% CI [-0.37, -0.02]). Finally, we assessed whether target inconsistency

similarly predicted participants' moral impressions of other unrelated people, controlling for Part

1 impressions. Though this relationship trended in the expected direction, the correlation was not

significant (β=-0.15, 95% CI [-0.35, 0.05]).

Taken together, character-violations – which occur for most people and involve a variety

of targets – predict perceivers' feeling that the world is a disorderly place and doubting of their

ability to judge character accurately. These events were not, however, associated with worsened

moral impressions of other people.

Table 2

Regression Predicting Each Dependent Variable (Meaning Violation, Confidence Judging

Character, Moral Impressions of Others) in Study 1

	Meaning Violation				**Confidence Judging Character**				**Moral Impressions of Others**			
Predictors	β	*SE*	95% CI	*p*	β	*SE*	95% CI	*p*	β	*SE*	95% CI	*p*
(Intercept)	**2.88**	**0.06**	**[2.76, 3.00]**	**<.001**	**3.76**	**0.09**	**[3.60, 3.93]**	**<.001**	**4.32**	**0.10**	**[4.12, 4.51]**	**<.001**
Inconsistency	**0.15**	**0.06**	**[0.03, 0.26]**	**.013**	**-0.19**	**0.09**	**[-0.37, -0.02]**	**.029**	-0.15	0.10	[-0.35, 0.05]	.144
Time 1 score[a]					**0.24**	**0.09**	**[0.06, 0.41]**	**.008**	**0.31**	**0.10**	**[0.11, 0.50]**	**.003**
R^2			.014				.026				.028	
N			373				373				373	

Note. Bolding indicates $p \leq .05$. *SE* refers to standard error. 95% CI refers to 95% confidence

interval. Continuous predictors are standardized.

[a]Time 1 score refers to the before-violation Time 1 measure of the relevant dependent variable

(Confidence Judging Character, or Moral Impressions of Others).

Study 2

Study 2 sought to directly replicate and expand on these findings. See pre-registration for

exclusion procedures, predictions, and methods

(https://osf.io/dzj3n/?view_only=f0f70652fe344e81b7d278c4ae85bb35). Unlike with Study 1,

we sampled equally for good-to-bad character-violations involving five major target types

(family members, romantic partners, friends, coworkers, and public figures), and assessed

different situational and dispositional factors surrounding character-violations (perceivers'

closeness to perpetrators, status as direct victims or observers, and dispositional preference for

closure and personality stability). We hypothesized that being direct victims and having higher

chronic preferences for closure and stability would be associated with worse recalled outcomes;

we had no directional hypothesis for the role of closeness. Finally, we changed the measurement

of key dependent variables such that participants reported how much experiences *violated* their

character judgments of other people (rather than simply reporting their judgments of others

around that time).

Method

Participants

We aimed for 320 participants per target type condition for 95% power to detect small

effects of d=0.30 between conditions following exclusions. We first recruited 2,908 MTurk U.S.

residents who had not completed Study 1 to complete the prescreening questionnaire; 2,510

passed a comprehension check testing understanding of instructions, and reported recalling

character-violations with one of the target types. These participants were then sent a separate

survey to complete for the main study; 1,983 agreed, and 1,624 participants completed the full

study. Of this, 1,582 participants were retained after passing one attention check and reporting

they read instructions and described a real experience (33% male, M=37 (SD=12.5) years old)

(sample per condition: coworkers: 316, friends: 314, romantic partners: 322, family members:

318, public figures: 312)).

Procedure

As part of the prescreening, participants completed the 15-item Need for Closure scale (NFC; Roets & Van Hiel, 2011; e.g., "I don't like situations that are uncertain," 1 ("Completely disagree") to 7 ("Completely agree")) and completed a 3-item measure of implicit theories of morality (Dweck et al., 1995; e.g., "A person's moral character is something basic about them, and it can't be changed very much," 1 ("Strongly disagree") to 7 ("Strongly agree")). All other continuous measures in the main study used scales from 1 ("not at all) to 7 ("extremely"). The main study included:

Part 1 (During-Violation)[1]. Participants indicated targets' age, gender, the impression-violating behavior, and victim status (whether they were directly victimized). As in Study 1, participants reported the inconsistency of the target's behavior and the meaning violation they felt. They next indicated how much these experiences directly violated their moral impressions of others in general (3-item, e.g., "How much did this experience violate your sense that most people are good people?") and violated their confidence judging character (3-item, e.g., "How much did this experience violate your beliefs that, compared to other people, you are accurate in your judgments of others' inner character?").

Part 2 (Before-Violation). Participants were asked about their before-violation closeness (1-item: "How close did you feel to [This Person] at that time (when you still held a positive impression of [This Person]?"), and before-violation character judgments of targets (as in Study 1, Part 1).

Part 3 (After-Violation). Participants reported their character judgments of the target following the character-violation, and their compartmentalization of the target's immoral

---

[1] The order of sections Study 2 differed from the order in Study 1. See SOM for explanation of re-ordering.

behaviors (1-item: "To what extent do you think [this person's] negative qualities should be considered separately from [this person's] positive qualities?").

Results and Discussion

Analysis Strategy

We assessed how inconsistency predicted the dependent variables (Model 1), whether closeness with the target moderated associations (Model 2), and whether victim status (Model 3), and other trait and demographic variables (Model 4) added additional predictive value.[2] To account for variance from different target types, all main analyses were conducted within a multilevel framework with random intercepts for target type using the R package LMER (as pre-registered). Our pre-registration additionally specified other analyses not of interest to the main research questions; results for these are reported fully in the SOM. Continuous predictors are standardized, and categorical variables with two levels are dummy-coded (reference = 0). Means and correlations of all relevant variables are shown in Table 3.

---

[2] Our pre-registration mistakenly specified controlling for participants' baseline (Before-Violation) scores. Doing so reduces meaningful variation in how much character-violations predict dependent variables. We therefore exclude these covariates, but we report the results of controlling for them in the SOM. Results do not differ meaningfully.

Table 3

Means, Standard Deviations, and Correlations of All Relevant Variables in Study 2

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
Part 2 (Before-Violation Measures)														
1 Closeness	4.9	1.8												
2 Character Judgment of Target (α = .94)	5.5	1.2	0.30***											
Part 1 (During-Violation) Measures														
3 Victim Status	0.5	0.5	0.32***	0.01										
4 Inconsistency (α = .65)	5.9	1.2	0.17***	0.34***	0.03									
5 Meaning Violation (α = .88)	3.6	1.6	0.23***	0.02	0.25***	0.09***								
6 Violation to Confidence Judging Character (α = .95)	4.3	1.5	0.25***	0.12***	0.14***	0.13***	0.32***							
7 Violation to Moral Impressions of Others (α = .96)	3.9	1.7	0.20***	0.08**	0.14***	0.09***	0.47***	0.41***						
Part 3 (After-Violation) Measures														
8 Character Judgment of Target (α = .94)	2.6	1.5	0.08**	0.04	-0.07**	-0.13***	-0.19***	-0.06*	-0.09***					

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
9 Compartmentalization	3.1	1.6	0.10***	0.03	-0.08**	-0.08**	-0.06*	-0.05	-0.06*	0.52***				
Demographic and Trait Measures														
10 NFC (α = .88)	4.2	0.8	0.07**	0.09***	0.00	0.07**	0.14***	0.16***	0.16***	-0.05	0.02			
11 Entity Theory (α = .90)	3.6	1.5	0.02	-0.01	0.01	-0.03	0.08**	0.04	0.09***	-0.03	-0.01	0.15***		
12 Age During Experience	30.6	12.4	-0.05*	0.10***	-0.03	0.09***	0.02	-0.06*	-0.03	-0.00	-0.01	-0.04	0.05*	
13 Gender[a]	0.7	0.5	0.07**	0.09***	0.06*	0.05*	0.01	0.02	0.05	-0.05*	-0.03	0.16***	-0.00	0.05*

Note. *$p \leq .05$, **$p \leq .01$, ***$p \leq .001$.

[a] Correlations with gender exclude 9 participants who did not identify with either gender. Gender is coded (1 = Female, 0 = Male).
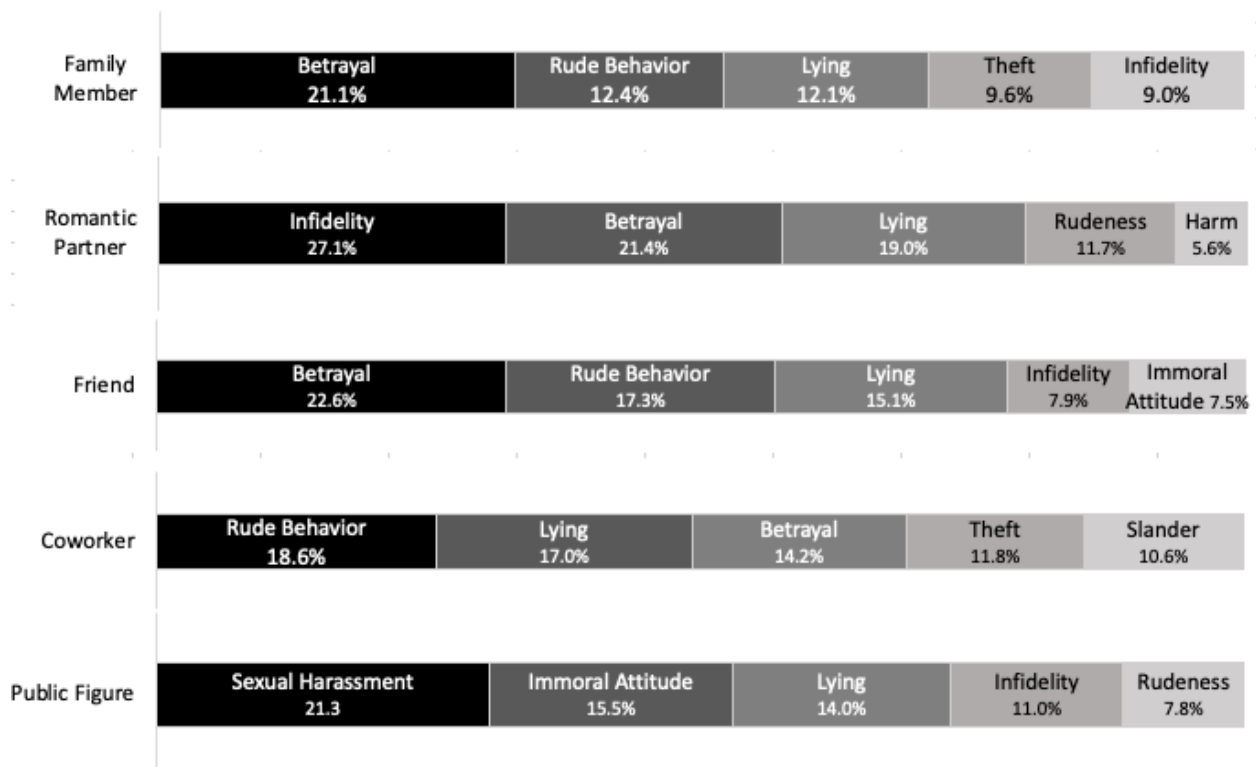
Victim Status is coded (1 = Yes, 0 = No).

Common Features of Character-Violations

Frequency and Descriptive Data. Almost all (94%) of the prescreened participants

recalled at least one character-violation experience. For participants in our main study, these

experiences most commonly took place within the first year of knowing the target. Nearly half

(45%) featured immoral behavior directed towards them. Targets were typically older than

participants, *M*=6.5 (*SD*=4.7) years older, and male (60%). In sum, these experiences again

appeared ubiquitous among participants, taking place with a wide range of targets, but usually

involving more men than women. Figure 2 shows common behaviors associated with each target

type.

Figure 2

Top Five Immoral Behaviors Listed for Each Target Type.

Family Member	Betrayal 21.1%	Rude Behavior 12.4%	Lying 12.1%	Theft 9.6%	Infidelity 9.0%
Romantic Partner	Infidelity 27.1%	Betrayal 21.4%	Lying 19.0%	Rudeness 11.7%	Harm 5.6%
Friend	Betrayal 22.6%	Rude Behavior 17.3%	Lying 15.1%	Infidelity 7.9%	Immoral Attitude 7.5%
Coworker	Rude Behavior 18.6%	Lying 17.0%	Betrayal 14.2%	Theft 11.8%	Slander 10.6%
Public Figure	Sexual Harassment 21.3	Immoral Attitude 15.5%	Lying 14.0%	Infidelity 11.0%	Rudeness 7.8%

Note. Percentages are out of the total sample size per target type condition.

Impression-Change. Participants reported dramatic changes in their judgments of targets' characters following character-violations (before-violation: *M*=5.5, *SD*=1.2, after-violation: *M*=2.6, *SD*=1.5), with little after-violation compartmentalization of targets' bad behaviors (*M*=3.1, *SD*=1.6). As expected, both compartmentalization and after-violation character judgments directly tracked perceived inconsistency (compartmentalization: β=-0.12, *SE*=0.40, 95% CI [-0.21, -0.04], *p*<.001; character judgment: β=-0.24, *SE*=0.04, 95% CI [-0.32, -0.17], *p*<.001).

Closeness with the target did not moderate these associations. Closeness did, however, independently predict more compartmentalization (β=0.19, *SE*=0.04, 95% CI [0.11, 0.27], *p*<.001) and more positive after-violation character judgments (β=0.13, *SE*=0.03, 95% CI [0.05, 0.20], *p*<.001). That is, participants who were closer to perpetrators were more likely to see their negative behaviors as separate from their positive qualities, and to maintain positive impressions of them.

Disruptions in Social Perception

Target Inconsistency. See Tables 4, 5, and 6 for full results with each dependent variable. In Model 1 analyses, we first assessed responses to the character-violations. Replicating Study 1, the inconsistency of a target's actions predicted greater meaning violation (β=0.15, 95% CI [0.07, 0.23]) and feeling less confident judging the character of others (β=0.20, 95% CI [0.13, 0.27]). Unlike in Study 1, we also found that inconsistency significantly predicted violation to participants' moral impressions of others (β=0.15, 95% CI [0.07, 0.23]). To the degree that targets violated perceivers' previous moral impressions of them, perceivers felt that the world was a less meaningful place, and they felt less confident and more negative in their views of other people's moral characters.

Closeness. In Model 2, closeness marginally moderated how inconsistency predicted meaning violation (β=0.07, 95% CI [-0.005, 0.14]), significantly moderated how inconsistency predicted violations to moral impressions of others (β=0.10, 95% CI [0.03, 0.18]), but did not moderate the relationship between inconsistency and confidence judging character (β=0.04, 95% CI [-0.03, 0.11]). That is, the closer participants felt to the target, the more the *same* inconsistent behaviors violated their sense of meaning and positive moral impressions of others. Closeness also had a sizeable independent effect of threatening confidence judging character (β=0.34, 95% CI [0.26, 0.42]). Taken together, closeness tends to exacerbate the impact of a character-violation.

We also re-analyzed Model 2 interactions with target type instead of closeness. We sum-coded target type (with the contr.sum function from the R package car), creating four sum-coded variables (each coding one of the target types as 1, three others as 0, and the reference group (romantic partners) as -1). This allowed us to use regressions to compare trends within each target type category to the average trend found across all target types. Unlike the analyses with closeness, we found few cases where findings with one target type differed from average findings. Only in the case of public figures did the relationship between inconsistency and decreased confidence judging character appear significantly weaker than expected (β=-0.16, 95% CI [-0.30, -0.02], *p*=.026). These findings suggest that character-violations with public figures may be unique, in that the outcomes may be less yoked with the degree of character-violation involved. We discuss these results more in the SOM.

Victim Status. In Model 3, we added victim status as a predictor, and found that it predicted more meaning violation (β=0.71, 95% CI [0.54, 0.88]), less confidence judging moral character (β=0.20, 95% CI [0.04, 0.36]), and more pessimism in moral impressions of others

(β=0.32, 95% CI [0.15, 0.49]). These effects were sizable and outweighed other factors,

suggesting that direct victimization was strongly associated with generalized doubt about order

in the world and the goodness of others.

Trait and Demographic Differences. In Model 4, participants' need for closure

consistently and independently predicted more meaning violation (β=0.20, 95% CI [0.12, 0.28]),

more threat to confidence judging character (β=0.21, 95% CI [0.14, 0.29]), and more negative

moral impressions of others (β=0.22, 95% CI [0.13, 0.30]). Participants who endorsed entity

theories of morality – beliefs about the stability of character – also experienced more meaning

violation (β=0.10, 95% CI [0.02, 0.17]) and more negative impressions of others (β=0.11, 95%

CI [0.03, 0.19]). They did not, however, differ in how much character-violations affected their

confidence judging others' moral characters.

Only one significant relationship emerged with demographic variables: The older

participants were, the less they felt that experiences threatened their confidence judging character

(β=-0.08, 95% CI [-0.15, -0.003]). With age, people reported being better able to maintain the

same level of confidence they had before these experiences. Notably, age was not associated

with other dependent variables.

Table 4

Multilevel Regression Predicting Meaning Violation from Predictors in Study 2

Predictors	Model 1				Model 2				Model 3				Model 4[a]			
	β	SE	95% CI	p	β	SE	95% CI	p	β	SE	95% CI	p	β	SE	95% CI	p
(Intercept)	3.62	0.13	[3.35, 3.88]	<.001	3.60	0.09	[3.42, 3.79]	<.001	3.29	0.12	[3.05, 3.52]	<.001	3.37	0.13	[3.12, 3.62]	<.001
Inconsistency	0.15	0.04	[0.07, 0.23]	<.001	0.11	0.04	[0.03, 0.19]	.010	0.10	0.04	[0.02, 0.18]	.011	0.10	0.04	[0.02, 0.18]	.018
Closeness					0.35	0.05	[0.26, 0.44]	<.001	0.28	0.05	[0.19, 0.38]	<.001	0.26	0.05	[0.17, 0.35]	<.001
Closeness X Inconsistency					0.07	0.04	[-0.00, 0.14]	.066	0.06	0.04	[-0.01, 0.13]	.116	0.06	0.04	[-0.02, 0.13]	.125
Victim Status									0.71	0.09	[0.54, 0.88]	<.001	0.70	0.09	[0.53, 0.87]	<.001
Need for Closure													0.20	0.04	[0.12, 0.28]	<.001
Entity Theory													0.10	0.04	[0.02, 0.17]	.014

	Model 1	Model 2	Model 3	Model 4[a]			
Age During Experience				0.06	0.04	[-0.02, 0.14]	.150
Gender				-0.11	0.08	[-0.28, 0.05]	.186
R^2	.017	.067	.073	.094			
N	1582	1580	1580	1560			

Note. Bolding indicates $p \leq .05$. *SE* refers to standard error. 95% CI refers to 95% confidence interval. Continuous predictors are standardized.

[a] Model 4, which assesses gender, excludes 9 participants who did not identify with either gender. Gender is coded (1 = Female, 0 = Male). Victim Status is coded (1 = Yes, 0 = No)

Table 5

Multilevel Regression Predicting Violation to Confidence Judging Character from Predictors in Study 2

	Model 1				Model 2				Model 3				Model 4[a]			
Predictors	β	*SE*	95% CI	*p*	β	*SE*	95% CI	*p*	β	*SE*	95% CI	*p*	β	*SE*	95% CI	*p*
(Intercept)	**4.34**	**0.12**	**[4.11, 4.58]**	**<.001**	**4.34**	**0.12**	**[4.11, 4.58]**	**<.001**	**4.34**	**0.12**	**[4.11, 4.58]**	**<.001**	**4.34**	**0.12**	**[4.11, 4.58]**	**<.001**
Inconsistency	**0.20**	**0.04**	**[0.13, 0.27]**	**<.001**	**0.20**	**0.04**	**[0.13, 0.27]**	**<.001**	**0.20**	**0.04**	**[0.13, 0.27]**	**<.001**	**0.20**	**0.04**	**[0.13, 0.27]**	**<.001**

	Model 1				Model 2				Model 3				Model 4[a]			
Closeness					**0.34**	**0.04**	**[0.26, 0.42]**	**<.001**	**0.32**	**0.04**	**[0.24, 0.40]**	**<.001**	**0.29**	**0.04**	**[0.21, 0.37]**	**<.001**
Closeness X Inconsistency					0.04	0.03	[-0.03, 0.11]	.247	0.04	0.03	[-0.03, 0.10]	.294	0.03	0.03	[-0.03, 0.10]	.317
Victim Status									**0.20**	**0.08**	**[0.04, 0.36]**	**.012**	**0.20**	**0.08**	**[0.05, 0.36]**	**.010**
Need for Closure													**0.21**	**0.04**	**[0.14, 0.29]**	**<.001**
Entity Theory													0.03	0.04	[-0.04, 0.11]	.372
Age During Experience													**-0.08**	**0.04**	**[-0.15, -0.00]**	**.041**
Gender													-0.08	0.08	[-0.23, 0.08]	.321
R^2	.017				.067				.073				.094			
N	1582				1580				1580				1560			

Note. Bolding indicates $p \leq .05$. *SE* refers to standard error. 95% CI refers to 95% confidence interval. Continuous predictors are standardized.

[a] Model 4, which assesses gender, excludes 9 participants who did not identify with either gender. Gender is coded (1 = Female, 0 = Male). Victim Status is coded (1 = Yes, 0 = No)

Table 6

Multilevel Regression Predicting Violation to Impressions of Other People from Predictors in Study 2

Predictors	Model 1				Model 2				Model 3				Model 4[a]			
	β	SE	95% CI	p	β	SE	95% CI	p	β	SE	95% CI	p	β	SE	95% CI	p
(Intercept)	3.92	0.06	[3.80, 4.04]	<.001	3.90	0.04	[3.82, 3.99]	<.001	3.76	0.08	[3.62, 3.91]	<.001	3.73	0.09	[3.56, 3.90]	<.001
Inconsistency	0.15	0.04	[0.07, 0.23]	<.001	0.12	0.04	[0.03, 0.20]	.006	0.11	0.04	[0.03, 0.19]	.009	0.11	0.04	[0.03, 0.19]	.010
Closeness					0.30	0.04	[0.22, 0.39]	<.001	0.29	0.05	[0.20, 0.38]	<.001	0.26	0.05	[0.17, 0.35]	<.001
Closeness X Inconsistency					0.10	0.04	[0.03, 0.18]	.009	0.10	0.04	[0.02, 0.17]	.012	0.10	0.04	[0.02, 0.17]	.012
Victim Status									0.32	0.09	[0.15, 0.49]	<.001	0.31	0.09	[0.14, 0.48]	<.001
Need for Closure													0.22	0.04	[0.13, 0.30]	<.001

	Model 1	Model 2	Model 3	Model 4[a]			
Entity Theory				**0.11**	**0.04**	**[0.03, 0.19]**	**.005**
Age During Experience				-0.04	0.04	[-0.12, 0.04]	.370
Gender				0.05	0.09	[-0.12, 0.22]	.589
R^2	.008	.044	.060	.082			
N	1581	1579	1579	1559			

Note. Bolding indicates $p \leq .05$. *SE* refers to standard error. 95% CI refers to 95% confidence interval. Continuous predictors are standardized.

[a] Model 4, which assesses gender, excludes 9 participants who did not identify with either gender. Gender is coded (1 = Female, 0 = Male). Victim Status is coded (1 = Yes, 0 = No)

In summary, Study 2 replicates core findings that recalled character-violations are associated with disruptions in people's experience of meaning and impressions of even unrelated others; findings also speak to additional factors that play a role. We explore additional analyses, including mediational pathways, in the SOM.

Study 3

Studies 1 and 2 found that participants associate remembered character-violations with negative psychological processes. However, these retrospective studies may capture faulty memories rather than actual events and outcomes. In Study 3, we aimed to experimentally replicate these core findings. We compared the effects of moral character-violations to a control: experiencing a consistently negative impression of a target's moral character. We hypothesized that a character-violation, in which perceivers encounter both good and bad information about a target, leads to greater consequences than encountering consistently bad information about a target. See pre-registration for exclusion procedures, predictions, and methods: https://osf.io/837r9/?view_only=68e1b4f4905d497c947961d523e73fdf.

Method

Participants

We sampled for 450 (225 per condition) U.S. residents from Prolific, giving us 95% power to detect an effect size of d=0.34 between conditions (our estimate was based on previous pilot testing). After passing attention checks, 446 remained (see SOM for full description; 41% Male, M=33.9 (SD=13.1) years old).

Procedure

Participants were asked to think of a setting from their own lives in which they meet new people. They then engaged in an impression-formation activity and were instructed to imagine that they had met a target in an everyday-life setting. At Time 1, participants in the moral character-violation condition were given positive character information about the target and told to imagine that he seemed "like a kind and likeable person," one who was "objectively caring, helpful, and upstanding". Conversely, participants in the control condition read negative character information; this target seemed "like an unkind and dislikeable person," one who was "objectively selfish, cruel, and hateful."

At Time 2, all participants read new information about the target engaging in immoral behavior, randomly reading one of two vignettes (depicting the target either engaging in hate speech or domestic abuse against a partner). A manipulation check followed: "To what degree would this new information change your impression of (Target Name)?", 1 ("Not at all") to 7 ("Extremely"). Participants then reported their sense of meaning violation (6-item; e.g., "To what extent do you think the occurrence of this experience would violate your sense that the world is a predictable place?" 1 ("Not at all") to 7 ("Extremely")). They reported threat to their confidence judging character (3-item; e.g. "This experience would make me strongly doubt whether I am a good judge of character in general"), and threat to their moral impressions of others (3-item; e.g. "This experience would change how good I think other people are in general"), both rated 1 ("Strongly Disagree") to 7 ("Strongly Agree"). Finally, they completed a demographic questionnaire.

Results and Discussion

Means and correlations of all relevant variables are summarized in Table 7.

Table 7

Means, Standard Deviations, and Correlations of All Relevant Variables in Study 3

Measures	*M*	*SD*	1	2	3
1 Impression-Change	5.0	2.2			
2 Meaning Violation (α = .90)	3.8	1.5	0.28***		
3 Threat to Confidence Judging Character (α = .96)	3.9	2.0	0.56***	0.47***	
4 Threat to Moral Impressions of Others (α = .93)	4.0	1.7	0.32***	0.59***	0.53***

Note. *p≤.05, **p≤.01, ***p≤.001.

Manipulation Checks

We first examined manipulation checks. As expected, participants in the moral character-violation condition reported greater impression-change from Time 1 to Time 2 (*M*=6.4, *SD*=1.1), compared to the control condition, *M*=3.6, *SD*=2.2; *b*=2.80, *SE*=0.16, 95% CI [2.48, 3.12], *p*<.001.

Consequences for Social Perception

Did the moral character-violation condition lead to a greater sense of meaning violation compared to the control condition? Yes – as shown in Table 8, participants who experienced a character-violation reported greater threat to meaning than those who experienced a confirmed negative character impression (*b*=0.64, 95% CI [0.37, 0.91]). Similarly, we found that the moral character-violation condition caused greater threat to participants' confidence judging character, compared to the control condition (*b*=2.45, 95% CI [2.15, 2.74]), and greater threat to participants' moral impressions of others, compared to the control condition (*b*=0.86, 95% CI [0.56, 1.16]).
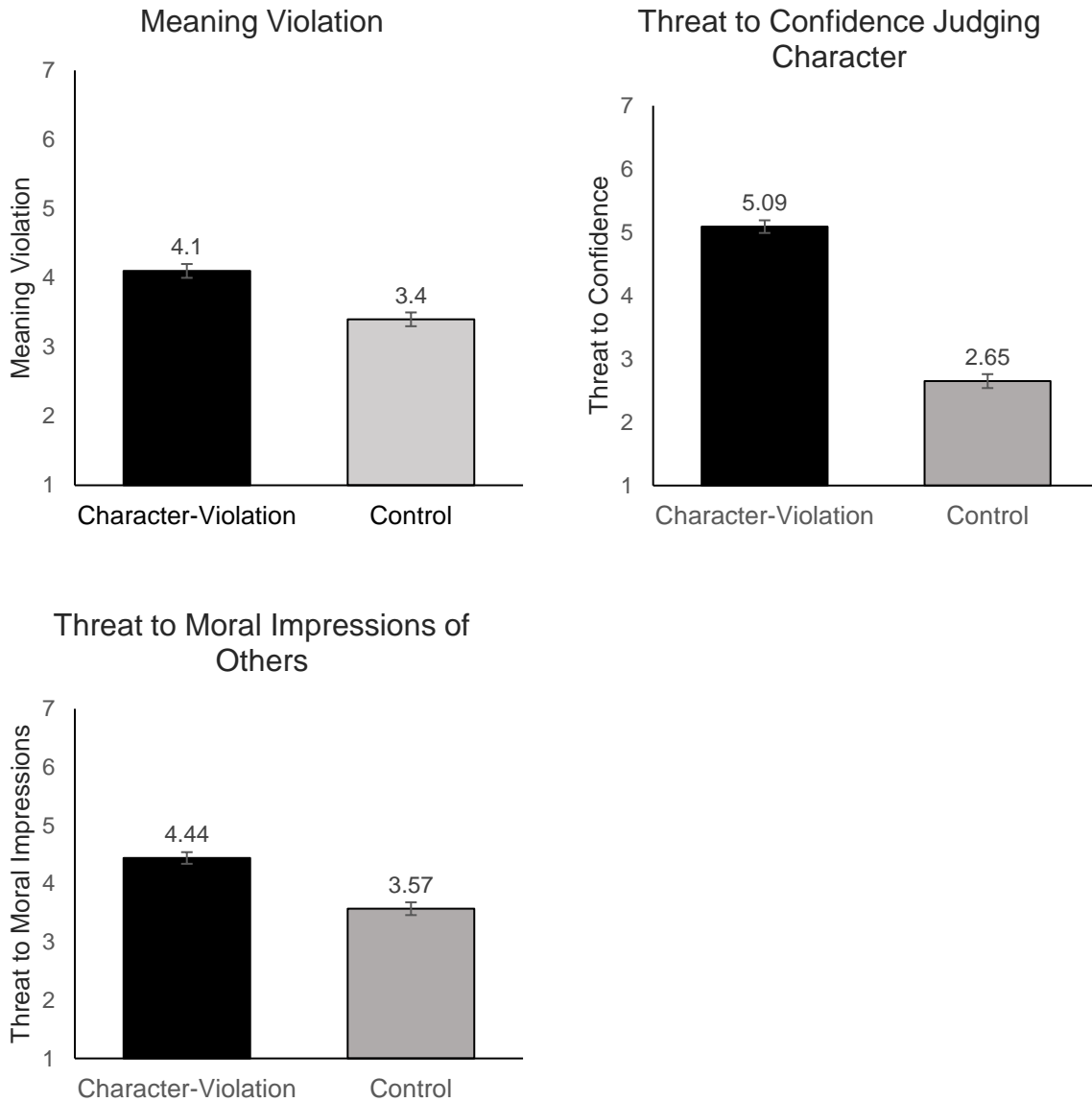
Experiencing a good-to-bad character-violation, in which a seemingly good target then appears bad, was across the board associated with greater consequences for perceivers' sense of meaning, confidence judging character in general, and expectations of others' morality, compared with having a negative character perception confirmed (see Figure 3). In other words, encountering an actor's consistent and prolonged immoral behavior affects perceivers less than does encountering a good individual who suddenly reveals immoral behavior. As with Study 2, we also explored additional mediational path models in the SOM.

Table 8

Regression Predicting Each Dependent Variable (Meaning Violation, Threat to Confidence Judging Character, Threat to Moral Impressions of Others) in Study 3

Predictors	Meaning Violation				Threat to Confidence Judging Character				Threat to Moral Impressions of Others			
	b	*SE*	95% CI	*p*	b	*SE*	95% CI	*p*	b	*SE*	95% CI	*p*
(Intercept)	**3.44**	**0.10**	**[3.25, 3.64]**	**<.001**	**2.65**	**0.11**	**[2.44, 2.86]**	**<.001**	**3.57**	**0.11**	**[3.36, 3.78]**	**<.001**
Moral character-violation condition	**0.64**	**0.14**	**[0.37, 0.91]**	**<.001**	**2.45**	**0.15**	**[2.15, 2.74]**	**<.001**	**0.86**	**0.15**	**[0.56, 1.16]**	**<.001**
R^2			.05				.37				.07	
N			446				446				446	

Note. Bolding indicates $p \leq .05$. *SE* refers to standard error. 95% CI refers to 95% confidence interval.

Figure 3

Means Per Condition in Study 3

Meaning Violation

Threat to Confidence Judging Character

Threat to Moral Impressions of Others

Note. Error bars represent standard error.

General Discussion

Moral character-violations appear frequently in the media and occasionally in everyday life. The present research provides an explanation of how these experiences affect perceivers, grounded in the meaning maintenance model (Heine et al., 2006) and social perception literature

(Goodwin et al., 2014). Across all three studies, good-to-bad character-violations were associated with disruptions in perceivers' sense that they understand the world, their confidence judging character, and their impressions of people's morality in general. In other words, the psychological impact is not restricted to people's views of specific character-violated targets, but spills over to color how people view *other people* more generally. Studies 1 and 2 also illuminated the types of moral character-violations people tend to encounter in everyday life, exploring additional situational and dispositional factors that predict stronger feelings of loss of meaning. Study 3 found causal evidence for these effects.

Our findings speak to general experiences, but a few key variables powerfully predict recalled outcomes. Directly victimized targets, and those with higher preferences for closure and personality stability reported greater disruptions in meaning. These findings line up with past evidence that being directly betrayed or transgressed upon leads to strong negative emotions (Adams & Inesi, 2016; Hutcherson & Gross, 2011), and having higher dispositional needs for stability and closure predicts more negative reactions to meaning-violations (e.g., Doherty, 1998).

Participants' closeness to perpetrators also affected predicted relationships. Though moral psychology rarely considers closeness (Bloom, 2011), recent research finds that closeness to perpetrators makes perceivers more lenient in their judgments (Hofmann et al., 2018; Linke, 2012; Waytz & Young, 2018). In Study 2, we found instead that closeness may play a paradoxical role; it predicted reports of more negative outcomes, but also less negative later impressions of targets. Participants who felt close to perpetrators may have been better able to assimilate the contradictory information into their previous positive impressions of targets, but closeness to these character-violating targets may have also exacerbated participants' sense of

uncertainty about their environments. Future research should investigate why closeness moderates some of these relationships but not others.

Beyond the correlational evidence, Study 3's experimental manipulations demonstrated that character-violations cause these aversive disruptions to people's mental models. These findings also revealed that good-to-bad character-violations affected participants more than bad character-confirmations. That is, having a positive character impression violated led to worse outcomes compared to what intuitively sees like a worse experience: encountering a consistently immoral person who demonstrates prolonged negative behaviors. This evidence of the unique impact of character-violations adds to existing findings on the importance of social cognition for people's need for order and prediction (Landau et al., 2004; Plaks et al., 2005).

Finally, the current findings provide a novel explanation for why revelations about people's perceived characters are so aversive: Moral character-violations threaten meaning, leaving people feeling endangered and less trusting of others. Such a broad reaction is likely adaptive, as erroneous moral judgments often result in real danger. Our theory and findings fit with new research showing people most value predictability in others (Jordan et al., 2017; Turpin et al., 2021; Walker et al., 2021); people may be instinctually avoidant of those whose moral character cannot be clearly judged, in order to avoid the material and epistemic threats associated with being wrong.

Limitations and Future Research

A primary limitation of our first two studies is their retrospective and correlational nature. Although this design is better posed to capture everyday experiences, it is vulnerable to alternative explanations regarding demand effects, retrospective bias, and alternative directional interpretations. To the extent this is true, our studies may better capture people's personal

theories about character-violations rather than actual experiences. However, Study 3's experimental replication of these core findings supports directional and causal interpretations. Future experimental research should further clarify the causal effects of the situational and dispositional factors surrounding these character-violations.

Though Study 3 provides experimental support for our theory, the study design also possesses limitations. Study 3 uses hypothetical vignettes that ask how participants would feel given an imagined event. This design carries the weakness of measuring participants' imagined rather than experiential responses to an actual event; however, the parallel findings with the recalled everyday-life events of Studies 1 and 2 partially allay these concerns. Future research can address this limitation by using longitudinal designs to measure participants' actual post-experience reactions to everyday character-violations.

A number of other questions remain: Are moral impression-violations unique in their effects? We speculate that impression-violations that go from bad to good, or concern other traits (e.g., competence) would be less aversive, given that people intuitively forgive past wrongs (Siegel et al., 2018) and weigh competence less in global impressions (Goodwin et al., 2014). Future research can further examine other types of impression-violations.

Conclusion

Though people care deeply about ascertaining others' moral characters, clearly all occasionally stand to encounter deep violations in their beliefs about who is good. The present findings advance a novel reason why people may bemoan such experiences: They threaten a core way in which we make sense of our social landscapes.

References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Adams, G. S., & Inesi, M. E. (2016). Impediments to forgiveness: Victim and transgressor attributions of intent and guilt. *Journal of Personality and Social Psychology*, *111*(6), 866–881. https://doi.org/10.1037/pspi0000070

Bloom, P. (2011). Family, community, trolley problems, and the crisis in moral psychology. *The Yale Review*, *99*(2), 26–43.

Brambilla, M., Sacchi, S., Pagliaro, S., & Ellemers, N. (2013). Morality and intergroup relations: Threats to safety and group image predict the desire to interact with outgroup and ingroup members. *Journal of Experimental Social Psychology*, *49*(5), 811–821. https://doi.org/10.1016/j.jesp.2013.04.005

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology*, *51*(1), 149–166. https://doi.org/10.1111/j.2044-8309.2010.02011.x

Christy, A. G., Kim, J., Vess, M., Schlegel, R. J., & Hicks, J. A. (2017). The reciprocal relationship between perceptions of moral goodness and knowledge of others' true selves. *Social Psychological and Personality Science*, *8*(8), 910–917. https://doi.org/10.1177/1948550617693061

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A

sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, *92*(2), 208–231. https://doi.org/10.1037/0022-3514.92.2.208

De Bruin, E. N. M., & Van Lange, P. A. M. (2000). What people look for in others: Influences of the perceiver and the perceived on information selection. *Personality and Social Psychology Bulletin*, *26*(2), 206–219. https://doi.org/10.1177/0146167200264007

De Freitas, J., Cikara, M., Grossmann, I., & Schlegel, R. (2017). Origins of the belief in good true selves. *Trends in Cognitive Sciences*, *21*(9), 634–636. https://doi.org/10.1016/j.tics.2017.05.009

Doherty, K. T. (1998). A mind of her own: Effects of need for closure and gender on reactions to nonconformity. *Sex Roles*, *38*(9), 801–819. https://doi.org/10.1023/A:1018825231362

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*(4), 681–735. https://doi.org/10.1017/S0140525X00032325

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, *6*(4), 267–285. https://doi.org/10.1207/s15327965pli0604

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*(5), 426–435. https://doi.org/10.1038/s41562-019-0590-x.Resolving

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–168. https://doi.org/10.1037/a0034726

Harmon-Jones, E., Amodio, D. M., & Harmon-Jones, C. (2009). Action-based model of

dissonance: A review, integration, and expansion of conceptions of cognitive conflict. In *Advances in experimental social psychology, Vol 41.* (pp. 119–166). Elsevier Academic Press.

Harris, L. T., & Fiske, S. T. (2010). Neural regions that underlie reinforcement learning are also active for social expectancy violations. *Social Neuroscience*, *5*(1), 76–91. https://doi.org/10.1080/17470910903135825

Hartley, A. G., Furr, R. M., Helzer, E. G., Jayawickreme, E., Velasquez, K. R., & Fleeson, W. (2016). Morality's centrality to liking, respecting, and understanding others. *Social Psychological and Personality Science*, *7*(7), 648–657. https://doi.org/10.1177/1948550616655359

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review*, *10*(2), 88–110. https://doi.org/10.1207/s15327957pspr1002_1

Heiphetz, L. (2019). Moral essentialism and generosity among children and adults. *Journal of Experimental Psychology: General*, *148*(12), 2077–2090. https://doi.org/10.1037/xge0000587

Hofmann, W., Brandt, M. J., Wisneski, D. C., Rockenbach, B., & Skitka, L. J. (2018). Moral punishment in everyday life. *Personality and Social Psychology Bulletin*, *44*(12), 1697–1711. https://doi.org/10.1177/0146167218775075

Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social-functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*(4), 719–737. https://doi.org/10.1037/a0022408

Janoff-Bulman, R. (1992). Shattered assumptions: Towards a new psychology of trauma. In

Shattered assumptions: Towards a new psychology of trauma. Free Press.

Janoff-Bulman, R. (2012). Meaning and morality: A natural coupling. *The Psychology of Meaning.*, 191–213. https://doi.org/10.1037/14040-010

Jordan, J. J., Sommers, R., Bloom, P., & Rand, D. G. (2017). Why do we hate hypocrites? Evidence for a theory of false signaling. *Psychological Science*, *28*(3), 356–368. https://doi.org/10.1177/0956797616685771

Koehler, J. J., & Gershoff, A. D. (2003). Betrayal aversion: When agents of protection become agents of harm. *Organizational Behavior and Human Decision Processes*, *90*(2), 244–261. https://doi.org/10.1016/S0749-5978(02)00518-6

Landau, M. J., Johns, M., Greenberg, J., Martens, A., Pyszczynski, T., Goldenberg, J. L., & Solomon, S. (2004). A function of form: Terror management and structuring the social world. *Journal of Personality and Social Psychology*, *87*(2), 190–210. https://doi.org/10.1037/0022-3514.87.2.190

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234–249. https://doi.org/10.1037/0022-3514.93.2.234

Lerner, M. J. (1980). The belief in a just world. In *The Belief in a Just World.* (pp. 9–30). Springer.

Levy, S. R., Chiu, C., & Hong, Y. (2006). Lay theories and intergroup relations. *Group Processes & Intergroup Relations*, *9*(1), 5–24. https://doi.org/10.1177/1368430206059855

Linke, L. H. (2012). Social closeness and decision making: Moral, attributive and emotional reactions to third party transgressions. *Current Psychology*, *31*, 291–312. https://doi.org/10.1007/s12144-012-9146-1

Park, C. L., Riley, K. E., George, L. S., Gutierrez, I. A., Hale, A. E., Cho, D., & Braun, T. D. (2016). Assessing disruptions in meaning: Development of the global meaning violation scale. *Cognitive Therapy and Research*, *40*(6), 831–846. https://doi.org/10.1007/s10608-016-9794-9

Piazza, J., Landy, J. F., & Goodwin, G. P. (2014). Cruel nature: Harmfulness as an important, overlooked dimension in judgments of moral standing. *Cognition*, *131*(1), 108–124. https://doi.org/10.1016/j.cognition.2013.12.013

Plaks, J. E., Grant, H., & Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception. *Journal of Personality and Social Psychology*, *88*(2), 245–262. https://doi.org/10.1037/0022-3514.88.2.245

Randles, D., Benjamin, R., Martens, J. P., & Heine, S. J. (2018). Searching for answers in an uncertain world: Meaning threats lead to increased working memory capacity. *PLoS ONE*, *13*(10), 1–22. https://doi.org/10.1371/journal.pone.0204640

Randles, D., Proulx, T., & Heine, S. J. (2011). Turn-frogs and careful-sweaters: Non-conscious perception of incongruous word pairings provokes fluid compensation. *Journal of Experimental Social Psychology*, *47*(1), 246–249. https://doi.org/10.1016/j.jesp.2010.07.020

Roets, A., & Van Hiel, A. (2011). Item selection and validation of a brief, 15-item version of the Need for Closure Scale. *Personality and Individual Differences*, *50*(1), 90–94. https://doi.org/10.1016/j.paid.2010.09.004

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. https://doi.org/10.1207/s15327957pspr0504_2

Siegel, J. Z., Mathys, C., Rutledge, R. B., & Crockett, M. J. (2018). Beliefs about bad people are

volatile. *Nature Human Behaviour*, *2*(10), 750–756. https://doi.org/10.1038/s41562-018-
0425-1

Skowronski, J. J., Carlston, D. E., & Hartnett, J. (2008). Spontaneous impressions derived from
observations of behavior: What a long, strange trip it's been (and it's not over yet). In *First
impressions.* (pp. 313–333). Guilford Publications.

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept
distinct from the self. *Perspectives on Psychological Science*, *12*(4), 551–560.
https://doi.org/10.1177/1745691616689495

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*(1), 159–171.
https://doi.org/10.1016/j.cognition.2013.12.005

Tomasello, M. (1999). The human adaptation for culture. *Annual Review of Anthropology*, *28*(1),
509–529.

Tullett, A. M., Teper, R., & Inzlicht, M. (2011). Confronting threats to meaning: A new
framework for understanding responses to unsettling events. *Perspectives on Psychological
Science*, *6*(5), 447–453. https://doi.org/10.1177/1745691611414588

Turpin, M. H., Walker, A. C., Fugelsang, J. A., Sorokowski, P., Grossmann, I., & Białek, M.
(2021). The search for predictable moral partners: Predictability and moral (character)
preferences. *Journal of Experimental Social Psychology*, *97*(July).
https://doi.org/10.1016/j.jesp.2021.104196

Walker, A. C., Turpin, M. H., Fugelsang, J. A., & Białek, M. (2021). Better the two devils you
know, than the one you don't: Predictability influences moral judgments of immoral actors.
Journal of Experimental Social Psychology, *97*(December 2020).
https://doi.org/10.1016/j.jesp.2021.104220

Waytz, A., & Young, L. (2018). Morality for us versus them. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology*. Guilford Publications. http://0-search.proquest.com.cisne.sim.ucm.es/docview/2010401048?accountid=14514%0Ahttps://ucm.on.worldcat.org/atoztitles/link?sid=ProQ:&issn=&volume=&issue=&title=Atlas+of+moral+psychology&spage=186&date=2018-01-01&atitle=Atlas+of+moral+psychology&au=Way

Weidman, A. C., Sowden, W. J., Berg, M. K., & Kross, E. (2020). Punish or protect? How close relationships shape responses to moral violations. *Personality and Social Psychology Bulletin*, *46*(5), 693–708. https://doi.org/10.1177/0146167219873485