


# Cognitive Blame Is Socially Shaped

Bertram F. Malle<sup>1</sup>, Steve Guglielmo<sup>2</sup>, John Voiklis<sup>3</sup>, and Andrew E. Monroe<sup>4</sup>

<sup>1</sup>Department of Cognitive, Linguistic, and Psychological Sciences, Brown University; <sup>2</sup>Department of Psychology, Macalester College; <sup>3</sup>Knology, New York, New York; and <sup>4</sup>Veris Insights, Washington, D.C.

Current Directions in Psychological Science  
2022, Vol. 31(2) 169–176  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/09637214211068845  
www.psychologicalscience.org/CDPS  


## Abstract

Blame is not only a cognitive process but also a social act of moral criticism. Such acts of criticism often serve to correct a transgressor's behavior but can be costly—to the moral critic, the transgressor, and the community. To limit these costs, blame is socially regulated: Communities set standards of evidence for blame and expect individuals to provide warrant, or justification, for their expressed judgments by pointing to appropriate evidence. We describe the path model of blame, which captures the cognitive processes that underlie blame judgments and that specify the kind of evidence that counts as warrant for blame. We then show how the varying costs of blaming put social pressure on the moral critic to be accurate and fair. We also identify conditions under which these pressures are weakened and standards of evidence decline: for example, when the transgressor has low status or is an out-group member, when the critic has high status or is anonymous, or when interactions are online. We close with numerous open questions that we hope will inspire further research into the hypothesis that cognitive blame is socially shaped.

## Keywords

cognitive process, judgment, moral psychology, social cognition, social influence

Blame is an inescapable part of social life. Its cognitive makeup is therefore likely to be shaped by its social role. However, psychological research on blame has often focused on its information-processing properties far more than on its social function. To understand blame, we contend, is to understand how blame works cognitively in light of what it achieves socially.

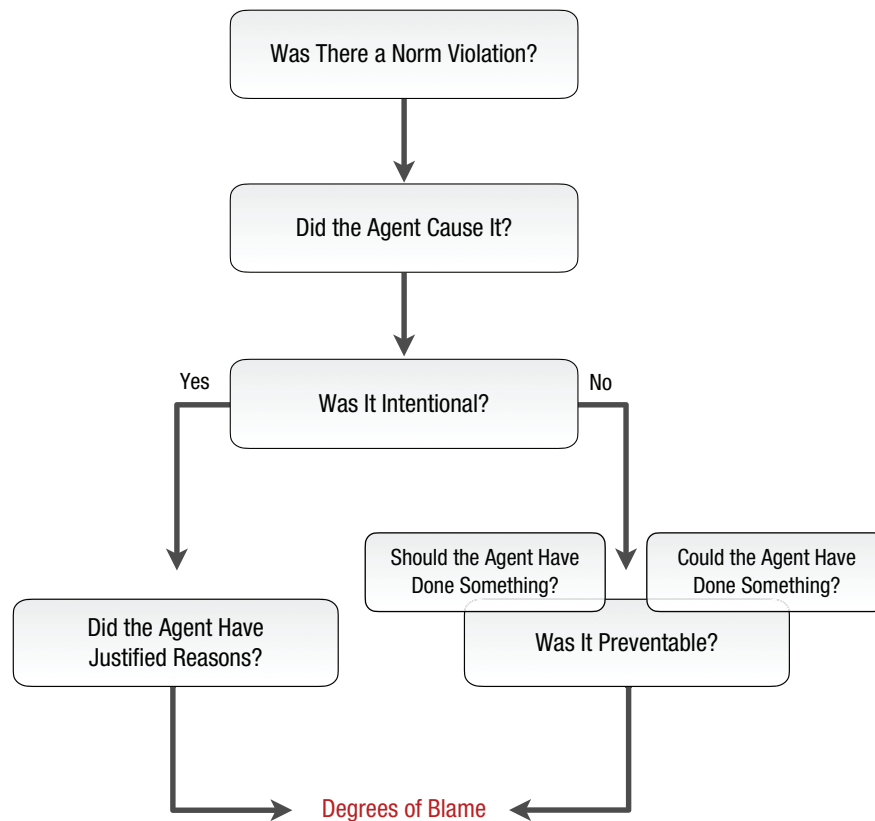
Blame has two sides. In public portrayals, it is primarily associated with conflict, accusation, and deflection. The “blame game” is an all-too-frequent reference in business and politics, and social media are rife with vicious acts of blaming. Psychological studies, too, have focused primarily on harmful victim blaming, self-blaming, and blame shifting. However, blame has another side: It may express rightful protest and a demand for justice (Ciurria, 2019), and considerable evidence has shown that judgments of blame are highly sensitive to information about the transgression and the transgressor in question, including causal, mental, and counterfactual information (Cushman, 2008; Monroe & Malle, 2019). How can these two sides of blame coexist? How can blame be so destructive, but also regenerative and sensitive to complex evidence? We suggest that this is because

the cognitive process of blame is socially regulated. When the community monitors and demands its fair use, blame is often prudent moral criticism; when the community leaves blame unchecked, it becomes a pernicious weapon of self-interest.

At its core, blame enforces the social and moral norms that guide human behavior—from expectations of politeness to prohibitions against harm. Norms constrain individual action in favor of benefits to the group, and without the enforcement of norms, many individuals would do what is in their own interest, to the detriment of the broader community. We propose that the cognitive structure of blame is an adaptation to the community's demands for fair norm enforcement—that is, enforcement grounded in evidence and proportional to the violation. Such fairness is important because norm enforcement is costly—certainly for the transgressor, but also for the blamer. The community therefore

## Corresponding Author:

Bertram F. Malle, Department of Cognitive, Linguistic, and Psychological Sciences, Brown University  
Email: bfmalle@brown.edu



**Fig. 1.** The path model of blame (adapted from Malle et al., 2014, Fig. 2; copyright 2014 by Taylor & Francis Ltd, www.tandfonline.com). The moral judge acquires information guided by an ordered set of questions whose answers yield the appropriate degree of blame.

sets standards of evidence and prescribes what information should be considered so that judgments of blame will be fair. Most moral judges will try to meet those demands, for blaming unfairly is itself a blameworthy act.

What, then, are the standards of evidence that dictate information processing en route to blame? Our path model of blame (Malle et al., 2014) proposes that people process the following information components (see Fig. 1): what norms the event in question violated; how causally involved the person was; whether that involvement was intentional; if it was intentional, what reasons the person had and how justified they were; and if it was not intentional, what the person should have done (obligation) and could have done (capacity) to prevent the violation. These information elements and their postulated ordering derive from ordinary social cognition. For example, asking about the agent's reasons makes sense only if the violation in question is believed to be intentional (Malle, 2004). Moral information processing

is therefore rooted in powerful and highly efficient cognitive mechanisms.

In what follows, we briefly review recent experimental findings supporting this model—findings that show how blame judgments operate when they are well regulated. Then we explain what this regulation consists of—how the costs of blaming raise standards of evidence—and under what conditions it breaks down. Thus, we show that our model of blame can account for both well-regulated and unregulated cases.

## Evidence in Support of the Model

### *Information acquisition*

If blame judgments have the ordered structure depicted by the path model, people should seek to acquire some information components (e.g., about causality) before other components (e.g., about intentionality). Results from three experiments supported these predictions

(Guglielmo & Malle, 2017). When people asked questions to acquire information about a negative event, their questions predominantly followed the sequence of the path model: People asked about causality (e.g., “Was the fire caused by a person?”), then intentionality (e.g., “Did he set the fire deliberately?”), and then either about reasons for intentional violations (e.g., “Did he want to get back at the neighbor?”) or about preventability for unintentional violations (e.g., “Was he being careless?”). This order remained stable when people selected information components from an array of options, and even under tight time pressure. However, this efficient information-processing order did not prevent people from flexibly integrating information by occasionally reconsidering an earlier component (e.g., the specific norms the behavior violated) in light of information about a later component (e.g., intentionality).

### **Updating blame judgments**

If people are motivated to render blame judgments that meet the community’s standard of evidence, and if the evidence the community prescribes for blaming includes such factors as intentionality, reasons, and preventability, then people should be sensitive to changes in this evidence and update their judgments. In six experiments (Monroe & Malle, 2019), we documented substantial and reliable blame updating for each of the information components in the path model, across different samples of participants and even when they were under cognitive load (i.e., performing a competing task). People showed no signs of confirmation bias (i.e., interpreting evidence to support their initial judgments), and they were as sensitive to evidence that mitigated blame as to evidence that amplified blame, thus displaying no default desire to blame that some authors have postulated (e.g., Alicke, 2000).

Additional studies have provided support specifically for the model’s path for unintentional violations. Martin and Cushman (2016) showed that blame markedly decreases when outcomes are uncontrollable (what our model refers to as unpreventable), and Niemi and Young (2016) showed that blame increases when perceivers believe outcomes to be preventable. Buckwalter and Turri (2015) further demonstrated that the two components of preventability—the capacity and the obligation to prevent harm—make independent contributions to blame.

The path model of blame and its supporting evidence appear to stand in opposition to findings by other researchers who proposed that blame often does not *result from* processing information about factors such as causality and intentionality but rather *precedes* and biases such information processing (Alicke, 2000;

Nadler & McDonnell, 2012). That is, in this view, people decide on blame first and then adapt their information processing to rationalize the predetermined blame. We set aside here our own interpretation of the extant evidence for such biased processing (e.g., Guglielmo, 2015; Malle et al., 2014) and instead introduce a broader perspective that can reconcile cases of careful information processing with cases of bias.

### **The Social Regulation of Blame**

Blame is not only assessed cognitively but also expressed socially. Acts of blame, reproach, and rebuke serve to criticize and (ideally) correct a transgressor’s behavior. However, such acts can be costly—they can damage the transgressor’s social standing, elicit retaliation against the blamer, and strain community relationships (Strimling & Eriksson, 2014). To limit these costs while still enabling effective norm enforcement, societies place normative demands on social expressions of blame (Friedman, 2013), including the social demand to provide *warrant* for blame. That is, people must be able to justify their expressions of blame to their community (Malle et al., 2014; Voiklis & Malle, 2018).

Currently, there is little experimental research on the role of warrant in blame specifically. However, research on punishment shows that when social communities view punishment as unwarranted, it is less effective at correcting social behavior (Herrmann et al., 2008). Conversely, warranted punishment increases cooperation (Kuwabara & Yu, 2017) and prevents norm-violating behavior (Xiao, 2013). To the extent that punishment and social blame are similar norm-enforcement mechanisms, these findings support the claim that perceived warrant for acts of blaming facilitates its corrective function.

In addition to facilitating norm enforcement, demand for warrant should facilitate accurate information processing on the part of the blamer (Monroe & Malle, 2019). Research has demonstrated that the goal to be accurate leads to more complex, less biased information processing (Kunda, 1990), and in these studies, factors that increase demand for warrant, such as expecting to justify one’s judgments or expecting to be evaluated by other people (Kruglanski & Freund, 1983; Tetlock, 1992), were directly manipulated. If demand for warrant is to be met, people must be able to access the kind of information that provides such warrant. Indeed, people are able to offer rich explanations of their blame judgments (Bucciarelli et al., 2008). These explanations differ from explanations of other moral judgments, containing precisely the information (e.g., on intentionality and reasons) that the path model predicts (Voiklis et al., 2016). Thus, the information processing of

**Table 1.** Costs of Acts of Blaming Incurred by Three Parties

Transgressor	Moral critic	Community
Bad feelings	Retaliation	Strife, damaged relationships
Damaged relationships	Damaged relationships	Potential loss of the
Loss of public standing	If the blame is unfair, loss of reputation and loss of community protection	transgressor's contributions If the blame is unfair, loss of in-group trust

cognitive blame delivers the kind of evidence that can meet the community's demand for warrant.

We have suggested that demands for warrant emanate from costs of blaming. Below we offer a closer analysis of the impact of these costs, which leads to a number of testable hypotheses. In developing these hypotheses, we treat changes in standards of evidence as the more general result of changes in costs of blaming, and we treat demand for warrant as one way in which the community marks strict standards. (Other ways include sanctions for failing to meet the standards, such as ostracizing individuals who make baseless accusations.) Under such demand for warrant, blamers search more thoroughly for relevant information rather than relying on prior assumptions, set a higher threshold for believing this information, and consider the target's actual reasons for acting and the actual evidence for preventability rather than making (typically uncharitable) assumptions. (For more details, see Malle et al., 2014, pp. 157–158.)

### Costs of Blaming and Standards of Evidence

Socially expressed blame, or acts of moral criticism, incur costs for the transgressor, the moral critic, and the community. Inaccurate blame carries additional costs,

especially for the moral critic. In this section, we consider how these costs (summarized in Table 1) guide the standards of evidence that the community imposes on the moral critic. We also identify factors that are likely to shift these standards by altering the costs incurred by the community or the moral critic (see Table 2).

#### *Common costs that guide standards of evidence*

For the transgressor, being the target of an expressed blame judgment often causes bad feelings, a threat to the relationship with the moral critic, possible loss of public standing, and potential subsequent punishment. Whether or not the critic takes the transgressor's costs into account, the community is likely to do so and put some pressure on the critic to heed standards of evidence; consequently, baseless accusations rarely fare well.

For the moral critic, a significant cost of expressing blame lies in potential retaliation from the transgressor (Balafoutas & Nikiforakis, 2012). Transgressors are particularly likely to retaliate if they regard the blame as excessive. Furthermore, when community members see the blame as unfair or inaccurate, they may support the transgressor's retaliation and socially reject the critic (Miller, 1999). Thus, if a moral critic expresses blame at all (rather than withholding it out of fear), such

**Table 2.** Selected Factors That Alter Costs of Moral Criticism for the Critic or Community and Thereby Change Standards of Evidence

Factors that increase costs and raise standards of evidence	Factors that decrease costs and lower standards of evidence
<i>High-status transgressors</i> <sup>a</sup> present a greater risk of retaliation and a greater potential loss of contributions to the community.	<i>Low-status transgressors</i> present little risk of retaliation or damaged relationships.
<i>Low-status critics</i> are vulnerable to retaliation, and their criticism may be dismissed.	<i>High-status critics</i> are protected against retaliation and accorded credibility by default.
<i>Critics bound to strong fairness norms</i> (e.g., a judge or teacher) commit a serious role violation when they issue unwarranted criticism.	<i>Anonymous critics</i> face no threat of retaliation and do not risk loss of reputation or relationships.
	<i>Groups that are not tight-knit</i> encompass relationships that are less valuable or easily replaceable, and community trust is already low.

<sup>a</sup>Under some conditions, high status can encourage harsher judgments (see the main text for discussion).

blame will normally have to meet strict standards of evidence so as to qualify for community support and protection.

In tight-knit communities of small to moderate size (e.g., neighborhoods, academic departments, villages), moral criticism and conflict among community members create a significant burden, and the whole community can suffer from the combination of wrongdoing, public censure, and damaged relationships (Allen, 2002). Further costs lie in the potential loss of community contributions from the censured individual. All these costs are even heavier in the case of unfair accusations. If the group permits such accusations, any member may be unfairly accused in the future, which undermines trust within the group. It is therefore in the community's interest to set strict standards that forestall such accusations.

### ***Factors that moderate costs and standards of evidence***

Additional factors can raise or lower standards of evidence, primarily because they alter the costs incurred by the moral critic and the community (see Table 2).

A transgressor's high status will often increase costs for the moral critic (who is likely to face retaliation; Rothschild & Miethe, 1999) and for the community (which may suffer a loss of valued contributions from the censured person). Critics of high-status transgressors are therefore forced to meet strict standards of evidence and may even be silenced or dismissed, while the privileged transgressor is protected from blame and accountability (Ciarria, 2019, Chapter 9). Under some conditions, however, "tall poppies are cut down," because the community jumps at the opportunity to criticize and sanction the transgressor, using loosened standards of evidence. This reverse pattern appears to stem from a community's anti-authoritarian leanings and their belief that the transgressor does not deserve the high status (Feather, 1994).

Low-status transgressors, by contrast, lack power and the opportunity to retaliate against a moral critic, and the community will be less concerned about damaged relationships or about losing the transgressor's community contributions. With both critic's and community's costs reduced, standards of evidence for blaming low-status transgressors will drop. For example, the mere suspicion that a Black man raped a White woman could rouse a lynch mob (Dray, 2002).

A moral critic with low social status is vulnerable to the transgressor's retaliation, and the community may dismiss the criticism (Ciarria, 2019). For example, women's statements of having been raped or children's statements of experiencing abuse can meet considerable skepticism

(Cheit, 2017). A moral critic with high status will be better protected from retaliation and enjoy greater credibility (Cooney, 1994). Thus, standards of evidence tend to be higher for critics with low status than for critics with high status. In rare cases, a high-power role may elicit increased standards of evidence; this is the case when the person's role inherently requires fair and accurate moral judgments. A department chair has to meet strict standards of evidence when accusing a faculty member of scientific misconduct, and if the evidence turns out to be false, the chair's fall from grace would be steep.

When moral critics are anonymous, such as in some online contexts, they express blame without confronting the violator, and—with little threat of retaliation or losing a valued relationship—they can get away with unfair and inaccurate moral judgments. Moreover, the community is not easily able to enforce standards of evidence when critics are anonymous, and such individuals are generally judged less negatively than people whose identities are known (Gino et al., 2010). Unfair moral judgments expressed by anonymous individuals are therefore more likely to be tolerated.

But even when moral critics are identifiable, blame is less regulated when expressed online, as powerfully analyzed by Wynn (see Gladstone, 2020). In online communities, or indeed any society of strangers, costs of blame are lower because the critic's community bonds are weaker and cannot easily be damaged. Community trust is also lower, and strife is common. Thus, without community solidarity, unfair blame remains unchecked, and people enjoy the spectacle of unfair accusations and resulting derogations (Bhargava, 2020).

### **When Regulation Breaks Down: The Intergroup Context**

We have proposed that the varying costs of blaming govern a community's standards of evidence, which encourage moral critics to achieve fair and accurate blame judgments. However, standards of evidence apply primarily to blaming the members of one's own community. If the transgressor is an out-group member, standards of evidence decrease, because community members will worry less about their damaged relationships with the transgressor or about loss of community trust, and they will care less about the costs imposed on the transgressor or simply assume that the person must have committed a violation worthy of sanctions (Rudert et al., 2018). Initial evidence shows that blame for out-group members is harsher and less sensitive to mitigating evidence than blame for in-group members (Monroe & Malle, 2019, Study 6), and out-group members tend to be sanctioned more severely than in-group members (Schiller et al., 2014).

When out-group members have equal standing with in-group members, they may demand warrant or potentially retaliate, but oppressed out-groups can do neither. People in power remain unchecked in their tendencies to favor stereotyping over individuating information. For example, immigrants are a frequent target of blame for crimes, despite evidence to the contrary (Light et al., 2020). And not too long ago, accusations that a Black individual committed a transgression remained unquestioned, required no evidence, and led to violence dressed as justice (Dray, 2002).

Weak standards of evidence also allow stereotypes and prejudice to bias blame judgments. Blamers may assume, without evidence, that the target had bad motives in violating a norm or had an obligation and capacity to prevent the violation; if these false assumptions remain unchallenged, inaccurate or harsh blame ensues (Niemi & Young, 2016). In such cases, the blamer may still process the assumed information in line with the path model (i.e., more blame for bad motives, more blame for preventable violations); but in the absence of demand to proffer actual evidence for the assumptions, blame is left unregulated and biased.

When intergroup bias occurs in online interactions, toxic situations emerge. People form online networks with ideologically similar others and direct their moral outrage toward ideologically dissimilar others (Crockett, 2017). These conditions foster loose standards of evidence, because people are rarely pressed to justify their judgments to an audience that already shares their values and beliefs, especially when the judgments are about out-group members.

### The Special Case of Close Relationships

Our analysis of costs of blaming applies to a variety of domains, such as workplaces, classrooms, or close relationships, and helps elucidate unique features of these domains. Consider blame in close relationships. The costs of blame to the transgressor will include moderately bad feelings and a temporarily disrupted relationship, but rarely a threat to the person's social reputation (unless the criticism occurs in public). The critic may also incur fewer costs than in other relationships, as retaliation is less likely. Costs to the community may be rare, too—unless children, who may be unable to exert demand for warrant, are involved. As a result of these lower costs in close relationships, expressed blame is more likely and less regulated. As a counterweight, inaccurate or unfair blaming in such relationships tends to be contested and may therefore be corrected (Dersley & Wootton, 2000).

However, in close relationships, too, a number of moderators increase costs and tighten regulation of

blame. If the relationship is at an early stage, conflict is more likely to end the relationship, so the costs of criticism, and standards of evidence, will be higher than in a more established relationship. If the transgressor holds significant power, the critic is at greater risk for facing retaliation and therefore will more tightly regulate moral criticism. Finally, if the expression of blame occurs in public, the transgressor's reputation will be at risk, the critic must moderate this expression in order to maintain a reputation of being fair and accurate, and the audience may incur costs of embarrassment or anxiety that the event will escalate.

### Open Questions

Evidence for the social regulation of cognitive blame is accumulating, but many elements of the model have yet to be tested. In particular, no studies on blame have included direct manipulation of demands for warrant, and none have measured variation in standards of evidence. We also do not know whether stricter standards of evidence elicit overall milder judgments or the kind of careful information integration that the path model predicts.

An intriguing question arises from the analysis of cultural variation in tightness-looseness (Gelfand et al., 2017). Do tighter cultures, which exhibit more, stronger, and clearer norms and enact more severe sanctions of norm violations, also have stricter standards of evidence for blame and other sanctions? On the one hand, because standards of evidence for blame are themselves norms (norms of norm regulation, or second-order norms; Eriksson et al., 2021), tighter cultures should have more, stronger, and clearer norms of blaming, and thus more severe sanctions for inaccurate and unfair moral criticism (which violates norms of blaming). On the other hand, if violations of all first-order norms (those that govern behaviors other than norm regulation) are sufficiently aversive to the community, then blame and sanctioning behavior should be loosely regulated so that no norm violator will be spared.

A cultural perspective also raises the question of how the social regulation of blame has emerged over time. Social regulation is unlikely to be an adaptation specific to morality because it can be found in numerous domains of cognition (Mercier, 2016; Tetlock, 1992). Rather, the regulation of blame follows a set of (second-order) norms that are learned, enforced, broken, and—as are other norms—context-specific (e.g., stronger for high-status members, weaker for out-group members). We believe that these norms are old, having emerged out of small-group living among nomadic hunter-gatherer societies in which costs of criticism and conflict are particularly high. Modern children are first socialized into similarly small

groups of intimates, but enacting the norms of blaming requires cognitive and communicative maturation over many years of socialization. It is not a trivial feat to fully integrate all evidence that informs blame and to track this information to provide warrant for one's judgment.

Finally, we have maintained that the breakdown of blame regulation invites harsher blame (e.g., on social media); but absence of blame regulation may sometimes soften blame. For example, if a community demands a severe judgment of a hated out-group member's transgression, then an individual blamer, if freed from these community demands (e.g., through online anonymity), may arrive at milder judgments. This, finally, raises a fundamental question: Are people, by default, responsible moral judges (who can be corrupted by social pressures), or are they fundamentally motivated by an urgent desire to blame (and must be regulated)? Whichever is closer to the truth, the social regulation of blame promises to shape humans to be generally more responsible, if imperfect, moral judges.

### Recommended Reading

- Ciurria, M. (2019). (See References). A thorough examination, with many real-world examples, of social forces that limit or dismiss oppressed groups' moral criticism of those in power.
- Friedman, M. (2013). (See References). A philosophical analysis of the costs of blaming and the resulting pressures for people to blame responsibly—that is, warranted by actual evidence of the facts at hand.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). (See References). The original formulation of the path model of blame and an initial analysis of the social forces that regulate blame.
- Tetlock, P. E. (1992). (See References). A comprehensive analysis and documentation of the impact of demands for explanation and justification in numerous domains of reasoning.
- Voiklis, J., & Malle, B. F. (2018). (See References). An analysis of the social dimensions of blame from the perspective of the path model, clarifying both how blame regulates social behavior and how social demands for warrant regulate blame.

### Transparency



Action Editor: Robert L. Goldstone

Editor: Robert L. Goldstone

#### Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

### ORCID iDs

Bertram F. Malle  <https://orcid.org/0000-0003-0845-9601>  
 Andrew E. Monroe  <https://orcid.org/0000-0002-9755-8726>

### Acknowledgments

We thank our collaborators over the years who have shaped our thinking about these topics, especially Corey Cusimano and Boyoung Kim, and we thank Emma Auerbach for her helpful comments on an earlier version of this article.

### References

- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*(4), 556–574. <https://doi.org/10.1037//0033-2909.126.4.556>
- Allen, D. S. (2002). *The world of Prometheus: The politics of punishing in democratic Athens*. Princeton University Press.
- Balafoutas, L., & Nikiforakis, N. (2012). Norm enforcement in the city: A natural field experiment. *European Economic Review*, *56*(8), 1773–1785. <https://doi.org/10.1016/j.euroecorev.2012.09.008>
- Bhargava, V. R. (2020). Firm responses to mass outrage: Technology, blame, and employment. *Journal of Business Ethics*, *163*(3), 379–400. <http://doi.org.revproxy.brown.edu/10.1007/s10551-018-4043-7>
- Bucciarelli, M., Khemlani, S., & Johnson-Laird, P. N. (2008). The psychology of moral reasoning. *Judgment and Decision Making*, *3*(2), 121–139.
- Buckwalter, W., & Turri, J. (2015). Inability and obligation in moral judgment. *PLOS ONE*, *10*(8), Article e0136589. <https://doi.org/10.1371/journal.pone.0136589>
- Cheit, R. E. (2017). A response to articles and commentaries on *The Witch-Hunt Narrative*. *Journal of Interpersonal Violence*, *32*(6), 1002–1023. <https://doi.org/10.1177/0886260516688889>
- Ciurria, M. (2019). *An intersectional feminist theory of moral responsibility*. Routledge.
- Cooney, M. (1994). Evidence as partisanship. *Law & Society Review*, *28*(4), 833–858. <https://doi.org/10.2307/3053999>
- Crockett, M. J. (2017). Moral outrage in the digital age. *Nature Human Behaviour*, *1*(11), 769–771. <https://doi.org/10.1038/s41562-017-0213-3>
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Dersley, I., & Wootton, A. (2000). Complaint sequences within antagonistic argument. *Research on Language and Social Interaction*, *33*(4), 375–406. [https://doi.org/10.1207/S15327973RLSI3304\\_02](https://doi.org/10.1207/S15327973RLSI3304_02)
- Dray, P. (2002). *At the hands of persons unknown: The lynching of black America*. Random House.
- Eriksson, K., Strimling, P., Gelfand, M., Wu, J., Abernathy, J., Akotia, C. S., Aldashev, A., Andersson, P. A., Andrighetto, G., Anum, A., Arian, G., Aycan, Z., Bagherian, F., Barrera, D., Basnight-Brown, D., Batkeyev, B., Balaus, A., Berezina, E., Björnstjerna, M., . . . Van Lange, P. A. M. (2021). Perceptions of the appropriate response to norm violation in 57 societies. *Nature Communications*, *12*(1), Article 1481. <https://doi.org/10.1038/s41467-021-21602-9>
- Feather, N. T. (1994). Attitudes toward high achievers and reaction to their fall: Theory and research concerning tall

- poppies. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 26, pp. 1–73). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60151-3](https://doi.org/10.1016/S0065-2601(08)60151-3)
- Friedman, M. (2013). How to blame people responsibly. *Journal of Value Inquiry*, 47(3), 271–284. <http://doi.org/revproxy.brown.edu/10.1007/s10790-013-9377-x>
- Gelfand, M. J., Harrington, J. R., & Jackson, J. C. (2017). The strength of social norms across human groups. *Perspectives on Psychological Science*, 12(5), 800–809. <https://doi.org/10.1177/1745691617708631>
- Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless + harmless = blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human Decision Processes*, 111(2), 93–101. <https://doi.org/10.1016/j.obhdp.2009.11.001>
- Gladstone, B. (Host). (2020, August 5). Making sense of “cancel culture” [Audio podcast episode]. In *On the media*. WNYC Studios. <https://www.wnycstudios.org/podcasts/otm/episodes/making-sense-cancel-culture>
- Guglielmo, S. (2015). Moral judgment as information processing: An integrative review. *Frontiers in Psychology*, 6, Article 1637. <https://doi.org/10.3389/fpsyg.2015.01637>
- Guglielmo, S., & Malle, B. F. (2017). Information-acquisition processes in moral judgments of blame. *Personality and Social Psychology Bulletin*, 43(7), 957–971. <https://doi.org/10.1177/0146167217702375>
- Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868), 1362–1367. <https://doi.org/10.1126/science.1153808>
- Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression formation, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19(5), 448–468. [https://doi.org/10.1016/0022-1031\(83\)90022-7](https://doi.org/10.1016/0022-1031(83)90022-7)
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- Kuwabara, K., & Yu, S. (2017). Costly punishment increases prosocial punishment by designated punishers: Power and legitimacy in public goods games. *Social Psychology Quarterly*, 80(2), 174–193. <https://doi.org/10.1177/0190272517703750>
- Light, M. T., He, J., & Robey, J. P. (2020). Comparing crime rates between undocumented immigrants, legal immigrants, and native-born US citizens in Texas. *Proceedings of the National Academy of Sciences, USA*, 117(51), 32340–32347. <https://doi.org/10.1073/pnas.2014704117>
- Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. MIT Press.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, 25(2), 147–186. <https://doi.org/10.1080/1047840X.2014.877340>
- Martin, J. W., & Cushman, F. (2016). Why we forgive what can't be controlled. *Cognition*, 147, 133–143. <https://doi.org/10.1016/j.cognition.2015.11.008>
- Mercier, H. (2016). The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences*, 20(9), 689–700. <https://doi.org/10.1016/j.tics.2016.07.001>
- Miller, W. I. (1999). In defense of revenge. In B. A. Hanawalt & D. Wallace (Eds.), *Medieval crime and social control* (pp. 70–89). University of Minnesota Press.
- Monroe, A. E., & Malle, B. F. (2019). People systematically update moral judgments of blame. *Journal of Personality and Social Psychology*, 116(2), 215–236. <https://doi.org/10.1037/pspa0000137>
- Nadler, J., & McDonnell, M.-H. (2012). Moral character, motive, and the psychology of blame. *Cornell Law Review*, 97, 255–304.
- Niemi, L., & Young, L. (2016). When and why we see victims as responsible: The impact of ideology on attitudes toward victims. *Personality and Social Psychology Bulletin*, 42(9), 1227–1242. <https://doi.org/10.1177/0146167216653933>
- Rothschild, J., & Miethe, T. D. (1999). Whistle-blower disclosures and management retaliation: The battle to control information about organization corruption. *Work and Occupations*, 26(1), 107–128. <https://doi.org/10.1177/0730888499026001006>
- Rudert, S. C., Sutter, D., Corrodi, V. C., & Greifeneder, R. (2018). Who's to blame? Dissimilarity as a cue in moral judgments of observed ostracism episodes. *Journal of Personality and Social Psychology*, 115(1), 31–53. <https://doi.org/10.1037/pspa0000122>
- Schiller, B., Baumgartner, T., & Knoch, D. (2014). Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. *Evolution and Human Behavior*, 35(3), 169–175. <https://doi.org/10.1016/j.evolhumbehav.2013.12.006>
- Strimling, P., & Eriksson, K. (2014). Regulating the regulation: Norms about punishment. In P. A. M. Van Lange, B. Rockenbach, & T. Yamagishi (Eds.), *Reward and punishment in social dilemmas* (pp. 52–69). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199300730.003.0004>
- Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331–376). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60287-7](https://doi.org/10.1016/S0065-2601(08)60287-7)
- Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 486–491). IEEE. <https://doi.org/10.1109/ROMAN.2016.7745207>
- Voiklis, J., & Malle, B. F. (2018). Moral cognition and its basis in social cognition and social regulation. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 108–120). Guilford Press.
- Xiao, E. (2013). Profit-seeking punishment corrupts norm obedience. *Games and Economic Behavior*, 77(1), 321–344. <https://doi.org/10.1016/j.geb.2012.10.010>