Abstract. Blame is both a moral judgment and an act of moral criticism of a person’s transgression. Such blaming can be biased or evidence-based, enraged or dispassionate, destructive or persuasive—in short, people can blame badly or they can blame well. Theories of blame have emphasized either the good side or the bad side of blaming, and this chapter will consider empirical evidence on both sides. But it provides a framework for organizing the conditions under which people are likely to blame well and the conditions under which they are likely to blame badly. The key insight is that good blame is characterized by systematic processing of evidence, and this processing is demanded by community pressures that try to minimize the costs of blaming. When these community pressures break down, distorting factors can corrupt the processing of evidence and lead people to be sloppy, incorrect, or unfair, hence blame badly.

Introduction
People express blame as an act of moral criticism of a person’s transgression. Such blaming can be biased or evidence-based, enraged or dispassionate, destructive or persuasive—in short, people can blame badly or they can blame well. In English, we have a rich vocabulary for all the shades in between, from criticize and reproach to slander and chew out. In one study (Voiklis et al., 2014), we found that people consider some acts of moral criticism (e.g., lash out, berate) to be far less socially acceptable than others (e.g., criticize, blame). But what makes an act of criticism socially good or bad? In that study, the more acceptable acts were also more thoughtful and less emotional, and they held more promise to continue a conversation with the transgressor. These correlates are informative, but they do not tell us exactly under what conditions people blame badly and under what conditions they blame well.

To understand these conditions of blaming badly or well we first need to be clear on what blame is. After clarifying the nature of blame, this chapter will explore the conditions of good and bad blame. I will describe the systematic processing of evidence that characterizes good blame, which is regulated by community pressures that demand careful information processing, with the aim of minimizing the costs of blaming. When these community pressures break down, distorting factors can corrupt the processing of evidence and lead people to be sloppy, incorrect, or unfair, hence blame badly.

What is Blame?
I begin with what blame is not. First, blame is distinct from other moral judgments (Cushman, 2008; Kneer & Machery, 2019; Malle, 2021), such as mere evaluation (considering some event “bad”), norm judgments (declaring something to be, for example, “inappropriate” or
“forbidden”), or moral wrongness judgments. Whereas moral wrongness classifies a (typically intentional) behavior as a fairly serious violation, blame applies to both intentional and unintentional violations and is directed at the transgressor. Moreover, blame lies on a fine-grained continuum that is highly sensitive to information about causal contributions, intentionality, justification, and more. Second, blame is also distinct from punishment (Baumard, 2011; Buckholtz et al., 2015; Malle, 2021), which is coercive, uncontestable, and irreversible (Malle, in press). Finally, blame is not just an emotion of anger or disgust, though these emotions can accompany blame. To define blame as “a negative affective response—characterized by other-condemning emotions” (Gill & Cerce, 2021, p. 1669) equivocates phenomena that are both empirically and theoretically distinct—and, ironically, emotion researchers have long considered blame itself to be one of the characteristic appraisals (and thus precursors) of anger (Scherer et al., 2001).

**Forms of Blame**

Blame comes in several forms. One is merely metaphorical. A storm may be blamed for uprooted trees and lives lost; but clearly we don’t walk up to the storm and morally criticize it. I also set aside another meaning, which is useful in some contexts: identifying merely who is to blame for a certain bad outcome. A terrible accident happens at a concert and people die, so we want to find out who is to blame: The organizers? The city that issued the permit? The police? This step is necessary when a negative outcome is caused by several people or causes, and we have to first single out those that deserve further scrutiny (“those responsible”). Once we have identified these appropriate targets of blame, the most interesting and complex form of blame arises: that of a continuous judgment of deserved moral criticism.

**Blame as Judgment—Blame as Action**

In her head, she deeply blames her boss for making fun of the intern’s lisp; but she does not say anything; she does not publicly express her moral criticism. Thus, blame can be a judgment (in the head) or an action (public expression). Whatever factors distort the judgment will, by implication, distort the act. However, the very fact that blame is publicly expressed creates pressures on the judgment to be correct and fair (or so I will argue). In this chapter we will shift back and forth between blame judgments and acts of criticism, because their very dynamic helps explain the conditions of blaming badly and blaming well.

**Blame is Normative**

In the case of a probability judgment or a causal judgment, there is a sensible way of assessing how well the judgment depicts some independent reality. But with blame, it makes little sense to ask whether it depicts an independent reality. We say that a transgressor “deserves” blame or “should” be blamed, which hints at a normative process. For example, the community prescribes whether certain kinds of evidence must be considered for blame (e.g., intentionality; Robbins et al., 2017) or whether a transgression is justified (e.g., violence to defend one’s honor; Cohen & Nisbett, 1994). Blame is thus governed by community norms for what it means to blame appropriately, or to blame well.
Blame Judgments — The Good

Blame judgments build on several other judgments—both moral and nonmoral (Cushman & Young, 2011; Malle, 2021). Once we know what kind of information processing normally underlies blame, we can ask what can go wrong, how people can stray from this processing and blame badly. My colleagues and I have proposed the Path model of blame to describe this characteristic information processing that generates blame (Malle et al., 2012; Malle, Guglielmo, et al., 2014).

The Path Model of Blame

As shown in Figure 1, the process begins when a perceiver detects an event (behavior or outcome) that violates one or more of the perceiver’s endorsed norms. This step is often overlooked and critical for several reasons. For one, the fact that the triggering event is a norm violation distinguishes blame judgments from judgments of other negative events. When a hurricane uproots trees and blows off roofs, people are distraught, but there is no blame to go around, because no social–moral norm has been violated. However, if afterwards the electricity doesn’t return for days and the rescue crews focus on the richer neighborhoods first, then norms have been violated and blame judgments are likely to arise. Second, norms vary widely across communities, so it should be no surprise that the same event is morally evaluated in different ways by different groups who have different norms—liberals and conservatives, young and old, and so on.

Given a norm-violating event, people look for why it occurred (Weiner, 1985). Causal judgments are difficult, even more so when the event involves multiple agents. People will put effort into this process but also make mistakes. Having no causal hypothesis is unsettling (Heider, 1958; Malle, 2004), and finding at least a preliminary one is pleasing (Gopnik, 1998). For the emergence of blame, it is necessary that at least one agent is identified as a (or the) cause of the norm-violating event.

If such an agent is identified, people want to know whether the agent caused the event intentionally. Intentionality judgments can be extremely fast (Decety & Cacioppo, 2012) or take longer (e.g., in criminal investigations), but they are critical for what happens next. For intentionality is not just an amplifier of blame but it bifurcates the moral perceiver’s information processing along two paths (for reaction time evidence, see Monroe & Malle, 2017).

Along the intentional path, people search for the agent’s reasons for the intentional norm violation. People search for reasons when explaining any intentional action (Buss, 1978; Malle, 1999, 2004), but norm-violating actions in particular. The inferred reasons influence blame by way of justifying (or not) the given norm violation. Justifications are reasons that not only explain why the person performed the action in question but make it “right” within a given norm system (Hermann, 2015). Somehow, the agent must persuade the moral critic that the chosen action was morally preferred over its alternative (Malle & Phillips, 2023; Wheeler & Laham, 2016). Good justifications can substantially lower blame (Monroe & Malle, 2019; Shaver, 1985), because even though the action violated some norms, it upheld more important norms (Gollan & Witte, 2008; Malle & Phillips, 2023).
Along the other path of blame processing, justifications are irrelevant, because there was no choice, no reasons. The agent caused a norm-violating event, and the event is contrasted with its counterfactual alternatives (Byrne, 2017): what the agent should have done (obligation) and could have done (capacity) to prevent it. If either one of these components is absent (the agent had no obligation or no capacity to prevent it), then blame is considerably reduced (Catellani et al., 2004; Martin & Cushman, 2016; Monroe & Malle, 2019).

This model has been criticized as being an idealization (Ciurria, 2019) and as ruling out the possibility that blame is distorted (Nadler, 2014). In one sense, the model is an idealization; it describes the paradigmatic, the “good” kind of blaming. Apparently, this is the kind of blaming we find under experimentally rigorous conditions, where blame judgments support the model—both in the past literature (see Malle et al. 2014, for a review) and in more recent studies that test specific predictions of the model (Guglielmo & Malle, 2017; Monroe & Malle, 2017, 2019). But there are other conditions under which people blame badly; but rather than rule out distortion, the Path model can help us determine when and how distortions occur (Malle, Monroe, et al., 2014, pp. 258–259). If the supportive experimental results emerge in conditions favoring good blame (Alicke, 2000; Ciurria, 2019), then we can reconcile these results with the undoubted existence of distorted blame by identifying both favorable and unfavorable conditions.

1 In the 2014 formulation of the model, we overlooked one additional element in this path: efforts to prevent. A person who should have and could have prevented an event is also expected to make efforts to prevent it. If the person did, blame will stay modest.
Favorable Conditions for Good Blame

Default Correctness Motivation

For cognition in general, a favorable condition for systematic information processing is sometimes called “accuracy motivation” (Kunda, 1990). Blame processing may have a similar motivation built in by default, but blame itself is not “accurate,” as there is no independent reality we can compare it to. Blame is better labeled “correct” (a more normative term) when it is in line with the community standards for what constitutes appropriate blame (Friedman, 2013). These standards, I have proposed, include: applying the relevant norms to the event in question; attempting to acquire valid information about the key components (who did what with what mental states, obligations, and capabilities), and expressing degrees of blame that are calibrated to the available evidence. There is evidence suggesting that people have default correctness motivation when forming blame judgments.

For one thing, sensitive and systematic information processing towards blame is neither affected by time pressure nor by cognitive load (Guglielmo & Malle, 2017; Monroe & Malle, 2019). This suggests that blame processing is cognitively efficient and normally does not require additional correctness motivation. Furthermore, the widely documented anchoring effect (that initial judgments, especially numeric ones, bias subsequent judgments) repeatedly failed to arise in Monroe and Malle’s (2019) blame updating results. Instead, people systematically adjusted their blame ratings in light of new information, regardless of their original blame rating. They also systematically adjusted their blame ratings no matter whether the new information favored increasing or decreasing blame. This symmetry in judgment updating contradicts the common claim about people having a “desire to blame” (Ames & Fiske, 2013; Mazzocco et al., 2004), which predicts that people should be more disposed to increase blame than to decrease blame. Together these results suggest that, as a default, people might be motivated to make correct blame judgments. Where would such built-in correctness motivation come from? The answer lies in the community’s handling of the costs of blaming.

Costs of Blaming

Directing moral criticism at a person imposes costs on that person, including feelings of hurt and threat, potential damage to the relationship, or loss of standing (Friedman, 2013; Malle et al., 2022). Most people are reluctant to impose costs on others (Crockett et al., 2014; Grossman, 1996) unless they feel they have a good reason, or warrant, to do so. Moral critics can incur costs, too, especially when they turn out to be wrong—that is, when the alleged transgressor did not deserve the blame, or the amount of blame, handed to them. Such costs for the blamer include threat of retaliation and possible loss of reputation. Finally, even the community carries costs when blame is assigned unfairly or excessively. Nobody wants to live in a community where people are disposed to haphazardly criticize and condemn one another.

All these costs build incentives for moral critics to make fair and correct (evidence-based) blame judgments. But how does the community check whether those judgments are indeed fair and correct?

Demand for Warrant

One mechanism to keep blame in check is by demanding warrant, which means that a moral critic must in principle be able to reveal the evidence on which the relevant blame judgment was based (Friedman, 2013; Malle et al., 2022; Malle, Guglielmo, et al., 2014;
McKenna, 2012; Voiklis & Malle, 2018). Demand for warrant is closely related to the notion of “accountability”—the expectation that one has to justify one’s judgments to an audience (Lerner & Tetlock, 1999). Researchers have found that accountability reliably limits bias in judgment tasks under three conditions (Lerner & Tetlock, 2003)—and warrant for blame meets all of them: The person is answerable to an unknown audience (the community at large); the judgment task requires no special training (blame certainly does not); and bias would emerge from lack of self-critical attention (e.g., overlooking evidence, relying on heuristics). Most researchers in the accountability literature assume that people normally do not feel accountable and are therefore subject to bias in their judgments; an accountability manipulation (e.g., asking people to justify their judgments) is needed to reduce that bias. Because studies that find systematic blame processing (e.g., Monroe & Malle, 2019) have no explicit manipulations of accountability built in, we may conclude that for most people the readiness to provide warrant has become internalized, or at least has been well practiced. People have learned that they would not get away with groundlessly blaming another person. How and why they have learned this practice can be explained by a second mechanism that can keep blame in check: the power of reputation.

Reputation

Humans have a strong need to belong to a community that values them (Baumeister & Leary, 1995). When this value is at stake, people show more cooperative social behavior (Wu et al., 2016). Normally, cooperation lies in complying with first-order norms, such as those of politeness or respect. But cooperation also lies in complying with second-order norms—the norms that govern appropriate norm enforcement (Strimling & Eriksson, 2014). If, as I have suggested, the costs of overly strong or unfair blaming are indeed high for both perpetrator and community, then such costly blaming will harm the blamer’s reputation—nobody will be popular for accusing and condemning countless community members. Thus, whenever a person publicly blames a transgressor, both the transgressor’s and the blamer’s reputation will be at stake, creating an incentive to blame correctly.

Despite all these forces that facilitate fair and correct blaming, distortions clearly do occur. I now turn to the unfavorable conditions under which such distortions emerge.

When and How Does Blame Go Wrong?

The ways in which something can go wrong are almost always more numerous and more varied than the ways in which it can go right. So we might consider a long list of states (e.g., anger), traits (e.g., prejudice), or processes (e.g., confirmation bias) that distort blame. But that will not be satisfying. Much of the literature on biased moral reasoning has already described numerous forms of distortion. What it has not done is explain them—it has not specified when and how biases occur. Some authors have proposed a general “desire to blame,” which (as I stated earlier) is incompatible with data from Monroe and Malle (2019), as well as with Mazzocco et al. (2004) and Nadler and McDonnell (2012), where mitigation (e.g., when critics are lenient on a person of outstanding character) is at least as strong as exacerbation. Even if a model of motivated reasoning allows for both mitigation and exacerbation, the details of how motives alter judgment must be specified. I will use the Path model of blame (2014), and its associated analysis of social regulation of blame (Malle et al., 2022; Voiklis & Malle, 2018), as a theoretical framework to better account for biased blame: when it occurs and how it occurs. This attempt will also try to answer critics who suggested that the Path model does not account for, or does not even allow for, biased blame judgments (Ciurria, 2019; Nadler, 2014).
When Does Bias Occur?

If the costs of blaming produce favorable conditions to be correct, then decreasing these costs will produce less favorable conditions (Malle et al., 2022). For the community and the critic, decreasing costs will weaken oversight, demand for warrant, and correctness motivation. Changes in the transgressor’s costs will figure into these equations only to the extent that critic and community perceive the transgressor’s costs to be lower (which will weaken diminishing their efforts to regulate) or higher (hence raising their efforts to regulate).

Decreasing Costs for the Critic

When the moral critic’s costs of incorrectly blaming decrease, we can expect fewer efforts to achieve correct blame. That is the case when people blame privately (in their head) or when they are anonymous; for in those cases, reputational concerns are absent (Kurzban et al., 2007). Likewise, when critics are protected from incurring costs because of their power position, there is less incentive to get blame judgments correct (Yin et al., 2022). Finally, in online communities, incorrect blame is less costly because, even when the critic is not anonymous, their relationship with the community is weak. An additional condition, which to my knowledge has not been tested, is that people may be less motivated to be correct when blaming very close others to their face (Malle et al., 2022). At least in long-lasting relationships (e.g., family, long-term partners), certain costs of blaming badly are less severe (e.g., loss of reputation) or less likely (e.g., end of the relationship). Such a pattern would explain why many illustrations of blaming badly come from close relationships (Furlong & Young, 1996; Pickard, 2013).

Decreasing Costs for the Community

When the alleged transgressor and the costs that the person accrues are of little value to the community, blame judgments will deteriorate. This is likely to be the case when an outgroup member is blamed. In this case, costs for the community are low because they do not lose out on the person’s contributions or their relationships with the person. As a result, there will be less demand on the critic to provide warrant and strive for correct blame. The literature on punishment has documented worse treatment of outgroup members (Van Assche et al., 2020). Likewise, a recent study across numerous countries (McKee et al., 2023) suggests that people consider norm violations as more morally wrong when perpetrated by a stranger than by a family member. By comparison, the data are surprisingly sparse for blame judgments in this intergroup context. Halabi et al. (2015) asked participants how much “responsibility” a target person had for a car accident whose cause was either unclear or likely caused by the person’s speeding. If the responsibility question elicited blame judgments, people mitigated their blame for the ingroup member when the causal contribution was unclear but withheld such mitigation for the outgroup member. Similarly, Monroe and Malle (2019) found that people granted less blame mitigation to outgroup members (people on the other side of the ideological spectrum) in conditions where such mitigation was granted to ingroup members, and they also blamed outgroup members slightly more across the board. These results highlight an ambiguity in the ingroup-outgroup differences for punishment and wrongness judgments. The differences may arise because people apply different norms to ingroup and outgroup; in this case, blame is not necessarily distorted. Alternatively, people may weigh unintentionality or justified reasons more strongly for ingroup members, hence mitigate blame more for the ingroup than for the outgroup. Such differential willingness to accept evidence would indeed be distortion.
The Reverse: Increasing Costs of Blaming

Occasionally the community cares very much about the transgressor, so their costs of blaming increase as they fear endangering their relationships with the person and losing the person’s contributions to the community (Malle et al., 2022). In such cases, communities put pressure on critics to get blame right, but sometimes so much so that critics face skepticism (Chiet, 2017) and sometimes backlash (Ciurria, 2020), especially when the critics were also the victim of the transgression. In a study by Kaiser and Miller (2003), white participants read about a black job candidate who was rejected and then voiced criticism of having been discriminated against (rather than explain it in some other way). The participants judged the critic as hypersensitive and a troublemaker, even in the case of blatant racism. So in these cases, the distortion does not lie in the critic but in the reception of the moral criticism by the community.

How Does Bias Occur?

Changes to Information Processing

On a popular view, moral judgments are flawed because they are “motivated.” That is, when reasoning about a norm violation, the moral judge prefers a particular conclusion and “this preference alters reasoning processes in a way that adjusts moral assessments in line with the desired conclusion” (Ditto et al., 2009, p. 312). The key claim is that the quality of people’s information processing suffers from the existence of a “preferred” judgment. When moral critics already know they want to blame a bad character, a hated group member, or the like, they engage in a process of “blame validation,” distorting or outright fabricating the evidence that warrants the desired amount of blame (Alicke, 2002).

Most moral communities, I had suggested, would not routinely accept blatantly fabricated or distorted evidence. Occasionally we may get away with fitting the facts to our preferred conclusion, but that becomes unlikely when our and the other person’s reputation are at stake and when others can challenge the distorted evidence. However, the question here is how do distorted judgments (perhaps less blatant ones) emerge when they emerge?

The Path model of blame offers a two-part answer. First, if people have to distort evidence, it will be just the kind of evidence that would normally be accepted as warrant for blame: having brought about a serious negative event in question, having done it intentionally, without justified reasons, or, if done unintentionally, in ways that could and should have been prevented. The research evidence that Alicke and colleagues have accumulated for distorted blame judgments typically take this form of distorted evidence for blame judgments, but still the familiar kind of evidence (see Guglielmo, 2015, for a detailed analysis).

Second, the distortion will have to be different for different information components. Perhaps the widest room for distortion lies in the perception of the initial piece of evidence: the person’s behavior. The objectively same physical movements through space can be perceived as aggressive or not (Kahan et al., 2012), as implicit consent or not (Kahan, 2010). It is slightly more difficult to portray an ambiguous behavior as intentional, which may require a supply of evidence showing that the person performed equally intentional actions in the past. Whether a person’s reasons are justified or not depends on relevant norms but also whether we believe the person really had those reasons. Fabricating preventability requires plausible counterfactuals of what the person could have done differently in the given context or convincing claims about obligations the person had to take preventive efforts. Thus, if the Path model is correct,
potentially distorted judgments are still likely to respect the kind of evidence that could in principle be offered as warrant for blame, but the evidence itself is weak, inaccurate, or made up.

The considerable cognitive efforts of the motivated, actively distorting reasoner may seem somewhat implausible. Blaming badly, it appears, is more often based on sloppy cognitive efforts, not on sophisticated manufacturing of evidence. Indeed, when there is little community oversight, and little internalized correctness motivation, the moral judge may go another route: by simply assuming the presence (or absence) of some of the information component, such as intentionality, reasons, and so on.

The picture I suggest, then, is that the framework of evidence that guides blaming well also guides blaming badly—except that, in the latter case, people distort or cut short the evidence-honoring processing that the community normally demands. This picture allows both for distortions of overblaming (most often cited in the literature) for distorted underblaming. For when we want to mitigate blame for those we admire, we also find past evidence of justified reasons or unintentionality, or cast the situation as impossible to prevent.

I will now briefly discuss a number of previously examined factors that can distort blame and examine whether the proposed picture can accommodate them.

Candidate Factors that Distort Blame

Beliefs. Rather than acquire new information about a transgressor and the situation at hand, moral critics sometimes strongly rely on previous beliefs. Two information components that can be readily disrupted by preconceived beliefs are intentionality and preventability, which are beliefs about whether the transgressor had control or choice over the violation. Feelings of power increase the perception that others have such control, and power thereby increases blame (Koo et al., 2023; Yin et al., 2022). People who believe that they themselves have considerable control over behaviors and outcomes extend this assumption to others, therefore blaming them more for negligence (Murray et al., 2023; Weiss et al., 2021). And even just activating the concept of choice increases victim blaming (Savani et al., 2011).

Stereotypes and prejudice constitute a powerful type of belief, and they can have a substantial impact on blame judgments at virtually all steps of the path. Though blame is typically not the dependent variable of choice in the stereotype and prejudice literature, it seems uncontroversial to suggest that some people are convinced that members of particular groups have bad motives, act intentionally on those motives, and therefore deserve blame for any norm-violating actions. Rape myth beliefs, for example, include assumptions about women’s causal contributions to instances of rape and their obligations to prevent them (Malle, Monroe, et al., 2014, p. 254; Niemi & Young, 2014). Identifying such pre-existing beliefs offers a more detailed explanation of prejudice-based blame than to merely state that blame is biased by prejudice (Ciurria, 2019).

States. Neff and Karney (2004) found that wives of heterosexual couples formed less favorable blame judgments about their spouses the more stress they experienced. Though the authors did not report specific analyses, the impact of stress could easily run through perceptions of intentionality (“he did that on purpose”) and inferred reasons (“he just wants to hurt me”). Anger is an oft-cited candidate for a blame-distorting state. Anger is said to encourage heuristic processing and snap judgments and focuses on the most immediately available information. For example, when aggressive individuals are in an angry state, they tend to make hostile inferences about others’ intentions (Tiedens, 2001), which is likely to foster harsher blame judgments. Lerner et al. (1998) found that people induced to feel angry were more punitive toward negligent actors. When people were accountable for their judgments, however, punitiveness declined and
was no longer predicted by anger but instead by considerations of the actors’ volition vs. external constraints. Thus, anger may have simplified people’s information processing whereas accountability encouraged them to take into account potentially mitigating factors.

The relationship between anger and blame is more complex, however, than a simple distortion picture suggests (Ciurria, 2019). Anger can be a response to repeated transgressions by the same person and therefore support rather than distort an updated blame judgment. Anger can also be a response to repeated transgressions by different people—as in the case of microaggressions, where every single act (e.g., asking “Where are you really from?”) may appear mild, but the cumulative effect is powerful and the resulting anger is diagnostic of a larger collective pattern of transgressions. Finally, anger can be the appropriate expression of a transgression’s seriousness; without some individuals’ outrage, the community might not acknowledge this seriousness.

**Character.** Some authors have argued that taking character into account when forming blame judgments is a distortion; the rational blamer should consider only the evidence directly relevant to the norm violation (Alicke, 2000). Within the framework of the Path model, people use character as input information from which they make critical inferences about intentionality, reasons, perhaps more. If we think of moral character as base rate information, it does not seem irrational to take base rates into account when making moral judgments in the real world. However, such inferences are justified only when character carries some diagnosticity for the inferred components; yet people might be inclined to make such inferences even when character is irrelevant to the norm violation in question (Nadler & McDonnell, 2012). Whether a piece of information is irrelevant, however, is difficult to determine. When experimenters provide character information in between-subjects designs, participants assume the information is diagnostic and draw inferences, whether accurate or not (Royzman & Hagan, 2017). Indeed, character manipulations routinely affect causal-mental inferences along with blame judgments (Nadler & McDonnell, 2012), so once more, the entire framework of blame processing appears to be available, whether the pertinent inferences are distorted or not.

**Outcome.** A final factor that is sometimes treated as distorting blame is outcome (e.g., the amount of harm or damage caused; Kneer & Machery, 2019). In the classic outcome bias cases, two protagonists are described as thinking and doing the same things (e.g., driving drunk), but one causes a much worse outcome than the other (e.g., hitting a tree vs. a pedestrian). Outcome bias is said to hold when people (in between-subjects designs) blame the person causing the worse outcome more than the other person (Mazzocco et al., 2004). But why should outcomes be ignored? Causing a worse outcome is a worse norm violation; thus, more blame is warranted. From the Path model’s perspective, people are likely to both evaluate the underlying norm violation indicated by the outcome (e.g., a person died or a tree was damaged) and also use outcome information to infer certain information components, such as what the person could have or should have known or how much preventive effort the person exerted. Unless the outcomes are causally decoupled from the protagonist’s mind and behavior, providing different negative outcomes will lead people to generate different counterfactuals (Byrne, 2017); these counterfactuals, whether defensible or not, drive blame judgments. The evidence from studies that aimed to show biased blame from outcome information are again compatible with the claim that people process the familiar causal-mental components of blame (Guglielmo, 2015; Mazzocco et al., 2004).

In sum, a number of candidate factors can distort blame information processing, but how they do so appears to be in line with the Path model of blame: the processing, even when flawed,
runs through the information components that the model describes (Guglielmo, 2015). The distortions operate as assumptions (e.g., of intentionality) or as inferences from information believed to be diagnostic (e.g., from character to motive), and they occasionally weaken the information processing itself (e.g., accepting flimsy evidence). Thus, we have a framework in hand that tells us how people blame badly, complementing the proposal for when they blame badly (when costs are low and no demand for warrant exists).

**Conclusion**

In the précis to a BBC program on blame (BBC News World Service, 2019), we read: “When things go wrong, we crave something or someone to blame. It’s an emotional response found in nearly every culture.” I suggest that nearly everything in this claim is false. When things go wrong, humans first try to find out why, and if the bad event was caused by a human, then a search for information begins to determine if, and how much, the person deserves to be blamed. This search process is not necessarily emotional or the result of a craving; it is socially regulated by norms of blaming, which try to minimize the costs of blaming; such regulation, however, is plausibly culturally universal. The regulation can break down, however, and people can fall into blaming badly. Exactly how often they blame badly and how often they blame well is impossible to assess; but because we live in social communities that are regulated in this way, we know the difference, both in others and ourselves. So we all can, and know how to, blame better.
References


