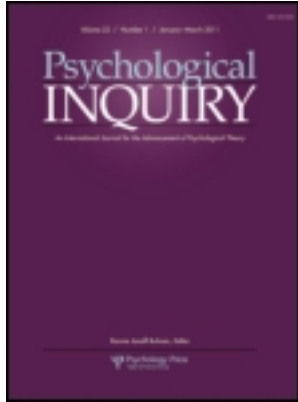


This article was downloaded by: [Geoff Goodwin]

On: 22 May 2014, At: 06:42

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Psychological Inquiry: An International Journal for the Advancement of Psychological Theory

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hpli20>

How Complete Is the Path Model of Blame?

Geoffrey P. Goodwin^a

^a Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania

Published online: 20 May 2014.

To cite this article: Geoffrey P. Goodwin (2014) How Complete Is the Path Model of Blame?, Psychological Inquiry: An International Journal for the Advancement of Psychological Theory, 25:2, 215-221, DOI: [10.1080/1047840X.2014.901138](https://doi.org/10.1080/1047840X.2014.901138)

To link to this article: <http://dx.doi.org/10.1080/1047840X.2014.901138>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

How Complete Is the Path Model of Blame?

Geoffrey P. Goodwin

Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania

Malle, Guglielmo, and Monroe's (this issue) theory of blame is a major intellectual and theoretical achievement. The model primarily focuses on the cognitive act of blaming by postulating a detailed series of processing steps that lead individuals to arrive at a judgment of blame (the Path Model of Blame). In doing so, the Path Model offers an elegant framework for conceiving of a disparate series of findings in the moral judgment literature. The model makes conceptual sense and accords intuitively with ordinary blame judgments. The postulated bifurcation of processing depending on whether an act is intentional or unintentional is a key insight. The conception of blame as depending on a series of processing steps brings a level of detail and granularity to the mental processes underlying blame that is welcome, and rare in current theorizing on moral judgment. In addition, rather than conceiving of blame as primarily biased, irrational, motivated, or disordered, the model highlights that most of the time, blame unfolds in a largely sensible and orderly fashion, if not always being entirely rational or impartial. The theory also draws our attention to important questions regarding social blame, an understudied topic. Finally, as the authors discuss, the model generates testable predictions, many of which have been supported in recent work that the authors review.

I am, therefore, a fan of the Path Model, and of the general approach to moral cognition that it encapsulates. It represents an important and substantial continuation of earlier theoretical remarks on blame made by these authors (Malle, Guglielmo, & Monroe, 2012).

Inevitably, however, some lingering questions exist with regard to the details and scope of the model. In this commentary, my goal is to highlight a series of questions and challenges that the model may need to face.

How Does the Model Take Into Account Transgression Severity?

A theory of blame ought to be able to account for the well-established effects of transgression severity on blame. One way that transgressions differ in their severity is in terms of their outcomes. Outcome information appears to influence blame even holding

constant the transgressor's intentions. For instance, Cushman (2008) showed that blame judgments are influenced by whether a harm occurred, or did not, despite transgressors' equivalent intentions. Other studies have revealed similar evidence for an "outcome bias" in judgments of blame, in which assessments of blame are distorted by chance factors that intervene between an agent's action and the eventual outcome, again holding constant the agent's intentions (e.g., Mazzocco, Alicke, & Davis, 2004; Shaver, 1970; Walster, 1966). Malle et al. are skeptical about the robustness of these effects of "outcome bias," in which outcomes—and individuals' corresponding assignments of blame—vary purely as a matter of luck.

However, transgressions also differ in their severity based on the magnitude of the harm that is intentionally inflicted. This, obviously, is also a form of outcome information, but one that is integral to the violation itself, and to the offender's intentions, rather than one that varies purely as a function of chance. A person's intending to cause a larger harm is clearly more blameworthy than a person's intending to cause a smaller harm, as many studies attest to. Robinson and Kurzban (2007) showed that ordinary individuals' rank orderings of the degree of blameworthiness of a variety of criminal offenses appears to take into account outcome severity information of this "integral" sort (see also Robinson, Goodwin, & Reising, 2010). Mental state information, such as an offender's degree of intentionality or motive, was also a key element underlying subjects' rankings of blameworthiness. But some key differences in Robinson and Kurzban's (2007) rankings primarily reflect the severity of the intended transgression or harm itself—intentionally stealing pies from a buffet was widely seen as less blameworthy than intentionally stealing a clock radio from a car, which in turn was seen as less blameworthy than intentionally stealing a microwave from a house. Similarly, intentionally slapping and bruising someone was widely seen as less blameworthy than intentionally punching someone, which in turn was seen as less blameworthy than intentionally stabbing someone to death (see also Carlsmith, Darley, & Robinson, 2002).

In these comparisons, causation, and intentionality information is held roughly (though clearly not perfectly) constant, with the degree of harm intentionally

caused being the primary driver of judgments of blameworthiness. These findings do not represent any sort of “bias” but rather a core aspect of the way blameworthiness is assigned. That is, in these cases, blame differs based on the magnitude of the harms caused, and on the *content* of the agent’s intentions—whether to cause larger or smaller harms—but not on the agent’s *degree of intentionality*, per se. It would therefore seem that a model of blame must somehow be able to account for these robust effects of the magnitude of the harm intentionally caused (transgression severity). However, it is not exactly clear how the Path Model accounts for effects of this sort, as the intentionality node in the model represents an assessment of the agent’s *degree of intentionality*, not an assessment of the *content* of the agent’s intentions. The model would therefore benefit from additional clarity in this respect.

One way to account for these effects of transgression severity would be to allow that, in the very first step of the Path Model—event detection—violations of norms are graded in terms of their seriousness. A violation of the norm not to steal property from others will be more severely or seriously violated when one steals a car, as opposed to a pencil. But this does not seem to be the approach that Malle et al. adopt. This first step of the Path Model represents an individual’s detection that a negative, norm-violating event has occurred. This process is regarded as being “very simple,” which suggests that, just as for other nodes of the model, it is also conceived in categorical *Yes/No* terms—did such a norm-violating event occur or not? Indeed, other aspects of the presentation of the model also suggest that the authors prefer to conceive of norm violations in more categorical terms: was a norm violated or not (*Yes/No*).

How, therefore, can Malle et al.’s model account for the difference in blameworthiness judgments as a function of *how bad* or *how negative* the event intentionally caused by the agent is; or alternatively, how important the norm violation is? One suggestion the authors make in the context of successful versus failed attempts is that successful attempts involve the violation of a greater range of norms (with each violation presumably being regarded as a separate binary instance): “The constitutive actions of trying to rob the bank usually violate fewer or weaker norms than the constitutive actions of actually robbing the bank” (p. 168). On this conception, the number of norms violated would presumably be summed or aggregated in some way to represent a final verdict of blame (although it would be useful to know more about how this summation process works, and by what the authors mean by “weaker” norms here). This approach might work when distinguishing successful attempts from failed attempts, as in the preceding example—a successful attempt at a crime will almost

invariably involve the violation of a greater range of norms than an unsuccessful one. It might also work to the extent that offenses can be subdivided into multiple component transgressions—stealing pies from a buffet involves only the act of theft, whereas stealing a microwave from a house involves the act of theft as well as the unwarranted entry into the house.

But, it is not clear that this approach can work for successful attempts that vary only in magnitude. Stealing a pencil does not seem to break a different norm than does stealing a car—rather, it breaks the same norm in a much more serious way (unless, of course, one were to posit separate norms for different acts of theft: don’t steal pencils; don’t steal cars; which would be psychologically unwieldy, and which would still require some additional weighting of the norms in terms of their seriousness).

Another option might be to package transgression severity information within the reasons part of the Path Model. A person’s reasons for stabbing someone to death are likely different from, and more culpable than, the reasons why they might slap and bruise another person. This idea has merit too, although it seems unlikely to be able completely to account for the effects of transgression severity on blame. There are easily conceivable pairs of cases in which two agents’ reasons and degree of intentionality are identical, with only harm or severity information varying (e.g., a person intentionally steals \$5 or \$50 from another person’s wallet, in both cases to help finance a consumer purchase). In cases of this sort, the contents of the agents’ intentions are clearly different—an intention to steal \$5 is distinct from an intention to steal \$50—but the agents’ intentionality, and their reasons are constant. Yet this difference in the severity of the violation (and of the content of the intentions) clearly influences blameworthiness. A second problem with relying on the reasons part of the model to account for transgression severity is that it seems unlikely that severity information would be processed so late (i.e., after causation and intentionality information is processed).

In sum, one way in which the Path Model could be improved is by including a more precise specification of how it can account for the severity of an agent’s violation in ordinary judgments of blame. It would also be helpful to know why the authors favor treating the event detection/norm violation part of the model as purely a categorical (*Yes/No*) matter, if this characterization is indeed correct. Perhaps one reason is that the authors are attentive to an important conceptual ambiguity with the notion of blame. Questions of blame seem to come in different guises. Sometimes we ask questions about blame that focus on issues of causal or moral responsibility: Who is to blame for this? Is he to blame? In these cases, questions of transgression severity are not relevant, but questions

of causation and intentionality are (reasons, obligation, and capacity information are also less relevant). And the natural response to such questions is often binary: Yes, he is to blame; No, he is not to blame. At other times, questions of blame seem to invite transgression severity information to be taken into account: How much blame does he deserve for this? How blameworthy are his actions? In these cases, the natural response is to apportion blame as matter of degree, in ways that reflect transgression severity, rather than in a strictly categorical matter. In these instances, arguably, causation information is partially presupposed, whereas outcome severity, intentionality, reasons, obligation, and capacity information all come to the fore. In this way, the notion of blame may not be free from some of the conceptual ambiguities that beset the notion of responsibility (as the authors note persuasively).

In What Ways Does the Model Allow for Graded Information Processing?

This ambiguity regarding graded versus categorical representation also emerges with respect to other parts of the model. In the presentation of the Path Model on page 151, most of its nodes are explicitly conceived of as binary *Yes/No* questions. Did an agent cause the negative event? Was it caused intentionally? If unintentional, did the agent have an obligation to prevent it? This categorical conception is reinforced by the authors' notion that a certain subjective threshold of evidence must be crossed before social perceivers assign causality, intentionality, and so on, seemingly in a binary fashion (p. 152). The evidence may be graded or probabilistic, but the blamer's representation ends up being categorical, one way or the other. The one clear exception concerns the agent's reasons for intentionally causing a negative event—for this question, the authors acknowledge that the quality of an agent's reasons is a matter of degree.

Clearly, this conception of these questions in *Yes/No* terms has the advantage of simplicity. Yet, as Malle et al. (this issue) also allude to at times, many if not all of these questions can also be seen as having graded answers. Causality may be seen as a matter of degree (p. 150), as might an agent's degree of intentionality, obligation, or capacity. Thus it is not entirely clear whether the *Yes/No* presentation of these questions in the model is simply a rhetorical device designed to render its basic theoretical insights more accessible, or whether it is instead meant to represent the core nature of the mental processes underlying blame. Allowing for the possibility of graded representations at each of these nodes might enhance the fidelity of the model.

Relatedly, it might be that the model could also benefit from allowing various "weights" to be attached to the different nodes, which would allow them to contribute more or less extensively to an overall judgment of blame. This might allow the model to capture individual difference in judgments of blame (e.g., some individuals may be particularly attentive to intentionality information in judging blame, whereas others focus more on reasons). And it may also prove to be a useful strategy in trying to account for the two different senses of blame just described. When a question of blame focuses on issues of causal responsibility—Who is to blame for this?—the agent causality node might be weighted more, whereas when a question of blame focuses on degree of culpability—How much blame does he deserve for this?—the additional processing nodes (intentionality, reasons, etc.) might receive more weight.

Can the Path Model Account for Motivational Effects of Blame on Lower Level Components?

The Path Model is a "blame late" model, in that blame "follows the logic of criteria" and is conceived as depending on a number of lower level assessments (see also Malle et al, 2012). In this way, it contrasts with "blame early" models, according to which blame is at least sometimes thought to occur rapidly and automatically, in a way that might bias later assessments of causation or intention (e.g., Alicke, 1992, who focuses on causation). At various points, the authors note that the Path Model might be able to accommodate some motivational aspects of blame, for example, "this focus on concepts and information processing in no way denies the role of affect and emotion in blame or the possibility of motivated reasoning" (p. 150); "arriving at a blame judgment is intertwined with emotion and motivation" (p. 152)—while also noting that the pathways specified represent only a default processing order, such that "the conceptual relationships allow for more flexible relations at the process level" (p. 152). Whether this means that they think the Path Model can account for findings showing that blame judgments may bias later judgments of causation or intention is not entirely clear. If so, it would be valuable to specify how the Path Model can accommodate such findings.

However, the fact that the authors go to considerable length to critique several of the most prominent "blame early" findings suggests that they do not take such findings to be compatible with their model. The authors' trenchant critique of these findings is

insightful, and they vividly portray the high hurdle that blame early models must overcome:

Vignette studies that try to demonstrate the undue effect of extra-evidential information face a nearly insurmountable challenge: Since people have to make judgments about ambiguous material, they are inferentially hyperactive and will inspect any information they receive for signs of what they want to know: the agent's causal role, mental states, obligations, preventive actions. Experiments without a ground truth will therefore have a difficult time making the normative distinction between justified and unjustified ("motivated") inferences. (p. 164)

Malle et al. (this issue) argue that this critique applies to many classic findings. As one illustration, they consider Alicke's (1992) famous example: A person speeding home to hide cocaine from his parents was judged more causally responsible for a later traffic accident he was involved in than a person speeding home to hide his parents' anniversary present. According to Malle et al., this result may be explained on the basis of rational inferences that "the drug-hiding agent was driving faster, was more inattentive, and more careless than the gift-hiding agent, warranting greater causality and blame judgments" (p. 163), rather than by the motivated account that Alicke proposes (an early blame judgment biases later judgments of causation).

This analysis is persuasive, but I am less confident that it can apply to all of the extant findings on motivated blaming in as convincing a way. In particular, a study by Nadler and McDonnell (2012, Study 2), which the authors discuss, provides a much more difficult challenge. In this study, the main protagonist, Sam Norton, stored oxygen tanks in his garden shed. When a group of youths dropped a cigarette alongside those sheds, the tanks exploded, resulting in the death of one of the youths. Sam Norton knew the tanks were a fire hazard. He had stored these tanks in his shed for one of three reasons: a laudatory reason (to provide medical care for his daughter who suffers from respiratory disease), a neutral reason (he is a businessman who provides in-home delivery of healthcare equipment), or a culpable reason (he is a football coach who illegally provides oxygen to his players). Malle et al. note that the fact that the subjects regarded Norton as most blameworthy in the case in which he stored the tanks to provide illegal oxygen to his players, and least blameworthy when he stored them to provide medical care, can be explained on entirely rational grounds which incorporate the agents' differentially culpable reasons en route to an eventual judgment of blame.

However, Malle et al. pass over the fact that the effect of a culpable motive was also observed on judgments of Norton's causal contribution to the

outcome, his intentionality in producing that outcome, and his responsibility for it. These results fit blame early models much better than they do the Path Model. And it is not at all obvious that these judgments can be rendered as being rational in the way needed by the Path Model, that is, in the way Malle et al. propose for Alicke's (1992) finding. The act of storing oxygen tanks is not a dynamic event, like driving a car, which may unfold in quite different ways depending on inferences about the actor's character. More important, whereas the cocaine-rescuing driver might plausibly be judged as more reckless or careless than the present-rescuing driver in Alicke's study, thus leading to a rational inference of his greater causal contribution, the negative character traits underlying the cheating football coach's behavior in Nadler and McDonnell (2012) are very different—they imply a shrewdly calculating, self-interested actor who is likely to take substantial precautions to ensure that his oxygen tanks are safe. A failure to do so would jeopardize his own livelihood and self-interest. It thus becomes hard to see why any rational grounds exist for regarding the culpable actor as more likely to have stored his tanks in a dangerous way, thereby making him more of a causal factor, or more intentional, in having caused the eventual explosion and death. In this way, this example seems precisely to fit the desiderata laid down by Malle et al. as comprising a stringent test of motivated moral judgment: The extraevidential information is such that "no diagnostic information (relevant to an interpretation of the norm violation) can be inferred" (p. 163) from it. This is therefore a case that cannot obviously be rendered consistent with the Path Model, instead providing support for blame early models (though admittedly, it would provide even more decisive evidence if participants' perceptions of the degree of caution Sam Norton took in storing his tanks had been measured and accounted for). In essence, the nature of the character traits underlying the actor's culpable motive in examples such as these matters. And this example counts as a reasonably convincing case in support of blame early processing, notwithstanding subjects' likely inferential hyperactivity.

What Sorts of Events Warrant Blame?

The Path Model is explicit that a negative, norm-violating event must be caused by an agent in order for blame to be warranted. However, several findings suggest that these criteria for blame might be overly stringent and that blame can be apportioned in a wider variety of ways. Agents may sometimes be blamed for purely mental events that do not cause any harm to anyone else, such as their hedonic reactions to their

violations, or their desire for harmful events to occur (that they do not cause to happen). And agents may also sometimes be blamed for events that do not violate any norms.

Some existing work has shown that actors' remorse or regret may affect moral condemnation and blame. When actors are remorseful, they are judged less severely and are more likely to be forgiven (Gold & Weiner, 2000; Schwartz, Kane, Joseph, & Tedeschi, 1978; Taylor & Kleinke, 1992). Criminals who do not display signs of regret or remorse may be more likely to receive the death penalty (Sundby, 1998). Conversely, actors who savor or enjoy their harm-doing are more likely to be seen as "evil," increasing observers' desire to punish them (Gromet, Goodwin, & Goodman, 2014; see also Ames & Johar, 2008; Szcurek, Monin, & Gross, 2012). Other results have also recently indicated that actors who possess "wicked desires," as a result of placing a bet that will pay off in the event of a tragic outcome, such as a hurricane, are blamed for these desires (Inbar, Pizarro, & Cushman, 2012).

In some of these cases, blame judgments were not directly assessed, which limits their immediate relevance to the Path Model. However, it seems plausible to assume that blame judgments would follow moral condemnation in most of these cases, even if not exactly in lockstep. If so, these results appear to point toward factors that lie outside the scope of the Path Model but that influence judgments of blame nonetheless—an actor's desire and an actor's hedonic states.

The Path Model may be able to accommodate some of these results without much trouble. An actor's remorse, or his pleasure, following a harmful act could provide diagnostic information about his original intentions, or his reasons for acting (see also Ames & Johar, 2009). And the nature of an actor's desires might also be thought relevant to his intentions and reasons for acting. And indeed, Malle et al. (this issue) consider cases in which agents may be blamed for failed attempts at harm, in which case the event that is blamed is the *plan* or *attempt* to cause harm (p. 147).

However, not all of these results are so easily accommodated, because they do not always involve any event beyond the actor's internal states, responses, or traits. For instance, Inbar et al. (2012)'s results suggest that an actor's wicked desires are seen as blameworthy in and of themselves. A man who places a bet on a hurricane occurring is only blamed for doing so when his eventual profit is contingent on the hurricane occurring. When the bet is simply a means of insurance (or hedging, such that his eventual profit is not contingent on the hurricane's occurrence), he is not blamed (Inbar et al., 2012, Study 3). Inbar et al. interpreted this as showing that the act of betting itself is not blamed—rather, it is the

"wicked desires" that the bet engenders (which are mutable depending on the contingency of the actor's profit) that are blamed. Similarly, in a study my colleagues and I have conducted, an observer who played no role in the causation of a harmful event was morally condemned for experiencing pleasure in the wake of the news of this event, although, admittedly, we did not directly assess blame (Gromet et al., 2014). Other work my colleagues and I have conducted showed that individuals are seen as especially responsible for their moral character traits (as compared with less morally relevant traits), though again, we did not directly assess blameworthiness here (Goodwin, Piazza, & Rozin, 2014).

Thus, although studies such as these have not always directly assessed blame, if it is indeed true that people are morally condemned and blamed for their desires, their hedonic reactions, and for their character traits, such results would cause some friction with the Path Model, because they show that judgments of blame can occur without any evidence of an actor's having caused a negative event. (This is, of course, unless the internal mental event, a desire, or a hedonic reaction, or the character trait, is seen as the "event" of consequence; but this does not seem to accord with an ordinary understanding of what an "event" is, and it raises difficult questions about whether the Path Model's subsequent criterion of intentionality can apply to such mental "events".)

Relatedly, there seem to be important cases in which a person is blamed for something that does not directly involve the violation of any norm. A football player might be blamed for causing his team to lose a game after fumbling a critical pass. A defender might be blamed for scoring an own goal. An employee might be blamed for transcribing critical financial information incorrectly. A researcher might be blamed for inadvertently failing to enter data correctly. In each case, the actor is blamed for an act of *incompetence* and not a moral failing involving the violation of a moral norm. To be sure, some cases of incompetence do seem to involve the violation of a norm—the employee or the researcher may be blamed for not having taken proper precautions to enter the critical information correctly and to check their work. Some cases of incompetence in sporting contexts may also be thought to reflect poorly on the actor's effort, preparation, or dedication, thereby indicating that the actor violated norms or expectations pertaining to these factors. But this does not seem true of all pure "accidents" of incompetence, which nonetheless seem to occasion blame—fumbling the football, or scoring an own goal—despite all the actor's efforts, best intentions, and precautions not to do so. Such accidents can arise in spite of great effort and dedication, and even in some cases, from too

much, rather than too little effort; and yet they still sometimes seem to occasion blame. Thus, these cases, although certainly not paradigmatic cases of blame, also seem to fall outside of the immediate purview of the Path Model.

What Counts as a Morally Eligible Agent?

The Path Model relies on the idea that only a morally eligible agent may be blamed. By this, Malle et al. (this issue) mean an agent who is capable of representing norms that may be violated. Accordingly, they treat a 12-year-old as a borderline case (pp. 153–154). It seems possible that this conception of a morally eligible agent, while eminently rational, is more restrictive than the notion that many ordinary individuals rely on in their actual blame judgments. In recent work my colleague Adam Benforado and I conducted, we found clear evidence for the existence of retributive motives directed toward animals that attack and kill humans, such as a shark that kills a swimmer (Goodwin & Benforado, in press). For instance, a shark that kills a more sympathetic victim is treated as more deserving of death, and more deserving of a painful death, than is a shark that kills a less sympathetic victim—in accordance with a retributive principle of proportional punishment (Kant, 1991). This result holds while statistically accounting for the perceived dangerousness of the respective sharks.

Showing that retributive motives—in low to moderate doses—can be directed toward animals is not the same as showing that such animals are morally blamed. But the idea of retribution does seem to presuppose some form of moral blame. Indeed, we found that a shark that had actually killed a human victim was regarded as far more blameworthy than one that had not (Goodwin & Benforado, in press, Study 4), and that blameworthiness ratings were close to the midpoint of a 9-point scale. Therefore, individuals regard these animal attackers as being at least moderately blameworthy for such attacks, and their desire to punish the animals significantly correlated with those blameworthiness assessments (other ongoing research we have conducted shows that blame and punishment for animals is correlated with the animal's presumed degree of intentionality, which fits with the Path Model). However, it seems unlikely that subjects would also regard these animal attackers as being capable of representing social norms, thus suggesting that this requirement of the Path Model is overly stringent. Blame, it seems, can be assigned to actors that do not have the capacity to represent norms. These findings suggest that the question as to whether the cause of a negative event

is a morally eligible agent may best be treated as a matter of degree, rather than as a categorical matter.

Is Constructive Moral Criticism Blame?

Malle et al.'s theory of blame distinguishes between cognitive blame (as depicted in the Path Model) and social blame. Important questions are yet to be answered regarding social blame, and the theory serves an important function by drawing our attention to these issues. One question, which the target article alludes to, is the form of moral criticism that is likely to be most beneficial, and most reformatory in everyday social life. The authors note, plausibly, that moral criticism (or blame) that is overly hostile, aggressive, or emotional is likely to prompt defensive maneuvers on the part of the targeted individual. In contrast, calmer, more rational, less aggressive moral criticism is likely to produce a more engaged response. But is this calmer, more rational form of criticism blame? It seems like of a stretch of the concept to regard this form of moral criticism as blame. (It is hard to imagine a marital therapist, for instance, advising one of the aggrieved parties in a marital dispute to engage in more blame, or even to engage in a different kind of blame. A more likely form of advice is to stop blaming altogether and to engage in an entirely different form of criticism.) The authors describe intriguing new evidence regarding the similarity between different forms of moral criticism and blame (e.g., Voiklis, Cusimano, & Malle, 2014), which seems to suggest that merely objecting to or disapproving of a particular form of conduct is not seen as similar to blame at all (see Figure 4 in Malle et. al, this issue); yet they also seem to want to regard calm, constructive moral criticism as a kind of blame. My question for the authors is therefore, Where does blame stop and these other forms of moral criticism begin?

A related question, which the authors do not mention but which I believe is important, is, How much social blame actually occurs in daily life? It strikes me that moral criticism, and even blame, is prevalent but that its expression usually takes the form of third-party gossip rather than direct moral reproach. Direct social blame seems to require a quite intimate connection between two parties, which serves as a buffer against the negative fallout that such blame can cause, or alternatively, a power differential between the blamer and the blamed, which also provides the blamer with sufficient warrant to blame. When two parties are equal in power, and not intimately connected, it is not obvious that blame will freely (or often) be attributed. Rather, it seems likely to be attributed cautiously, and only for rather serious social violations.

Concluding Comments

The theory of blame advanced in the target article is provocative, deep, and rich with insight, and it provides the most comprehensive existing theory of blame in the moral judgment literature. It is almost certainly right in its core theoretical tenets. The theory's emphasis on information processing is a welcome shift in the way moral psychologists often conceive of their endeavor and will provide a template for researchers in studying moral judgment in all of its many guises. I have raised critical comments in this commentary, which are intended to spur debate and new investigations. The criticisms I have raised do not challenge any of the core tenets of the theory but rather point to some aspects of blame that might fall outside the immediate purview of the theory, which it does not yet fully account for.

Note

Address correspondence to Geoff Goodwin, Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Solomon Lab Building, Philadelphia, PA 19104. E-mail: ggoodwin@psych.upenn.edu

Acknowledgments

I thank Dena Gromet and Ed Royzman for valuable comments on an earlier draft of this commentary.

References

- Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, 63, 368–378.
- Ames, D. R., & Johar, G. V. (2009). I'll know what you're like when I see how you feel: How and when affective displays influence behavior-based impressions. *Psychological Science*, 20, 586–593.
- Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish?: Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, 83, 284–299.
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108, 353–380.
- Gold, G. J., & Weiner, B. (2000). Remorse, confession, group identity, and expectancies about repeating a transgression. *Basic and Applied Social Psychology*, 22, 291–300.
- Goodwin, G. P., & Benforado, A. (in press). Judging the goring ox: Retribution directed towards animals. *Cognitive Science*.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106, 148–168.
- Gromet, D. M., Goodwin, G. P., & Goodman, R. A. (2014). *Pleasure at another's pain: The influence of a target's hedonic states on attributions of evil*. Manuscript in preparation.
- Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin*, 38, 52–62.
- Kant, I. (1991). The metaphysics of morals (H. B. Nisbet, Trans.). In H. Reiss (Ed.), *Kant: Political writings* (pp. 131–175). Cambridge, UK: Cambridge University Press. (Original work published 1797)
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2012). Moral, cognitive, and social: The nature of blame. In J. Forgas, K. Fiedler, & K. Sedikides (Eds.), *Social thinking and interpersonal behaviour* (14th Sydney Symposium of Social Psychology; pp. 313–341). Philadelphia, PA: Psychology Press.
- Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, 26, 131–146.
- Nadler, J., & McDonnell, M.-H. (2012). Moral character, motive, and the psychology of blame. *Cornell Law Review*, 97, 255–304.
- Robinson, P. H., Goodwin, G. P., & Reisig, M. D. (2010). The disutility of injustice. *New York University Law Review*, 85, 1940.
- Robinson, P. H., & Kurzban, R. (2007). Concordance and conflict in intuitions of justice. *Minnesota Law Review*, 91, 1829–1907.
- Schwartz, G. S., Kane, T. R., Joseph, J. M., & Tedeschi, J. T. (1978). The effects of post-transgression remorse on perceived aggression, attributions of intent, and level of punishment. *British Journal of Social and Clinical Psychology*, 17, 293–297.
- Shaver, K. G. (1970). Defensive attribution: Effects of severity and relevance on the responsibility assigned for an accident. *Journal of Personality and Social Psychology*, 14, 101–113.
- Sundby, S. E. (1998). Capital jury and absolution: The intersection of trial strategy remorse and the death penalty. *Cornell Law Review*, 83, 1557–1598.
- Szczurek, L., Monin, B., & Gross, J. J. (2012). The stranger effect: The rejection of affective deviants. *Psychological Science*, 23, 1105–1111.
- Taylor, C., & Kleinke, C. L. (1992). Effects of severity of accident, history of drunk driving, intent, and remorse on judgments of a drunk driver. *Journal of Applied Social Psychology*, 22, 1641–1655.
- Voiklis, J., Cusimano, C., & Malle, B. F. (2014, July). *A social-conceptual map of moral criticism*. Paper presented at the 36th Annual Meeting of the Cognitive Science Society, Quebec City, Canada.
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, 3, 73–79.