ARTICLE IN PRESS

CellPress
REVIEWS

Review

A Theory of Moral Praise

Rajen A. Anderson,[1,*] Molly J. Crockett,[2] and David A. Pizarro[1]

How do people judge whether someone deserves moral praise for their actions? In contrast to the large literature on moral blame, work on how people attribute praise has, until recently, been scarce. However, there is a growing body of recent work from a variety of subfields in psychology (including social, cognitive, developmental, and consumer) suggesting that moral praise is a fundamentally unique form of moral attribution and not simply the positive moral analogue of blame attributions. A functional perspective helps explain asymmetries in blame and praise: we propose that while blame is primarily for punishment and signaling one's moral character, praise is primarily for relationship building.

Moral Praise

A great deal of psychological research has focused on understanding the processes involved in moral attribution, describing when and how individuals are associated with and then held accountable for actions of moral importance. However, the most influential theories of moral attribution have focused almost exclusively on blame, the attribution of responsibility for an immoral act [1–4], with less attention applied to judgments of positive moral acts and the attribution of praise. However, as psychological research on moral praise accumulates, it is becoming apparent that praise and blame are far from attributional mirror-images of each other [5–9]. Moral praise appears to be a unique form of moral judgment.

By moral praise, we mean the cognitive appraisal regarding an agent's positive moral behavior and character (e.g., being prosocial, performing a 'good deed', helping others in need, etc.), the appraisal that an agent's positive behavior exceeded typical duties and obligations (what philosophers refer to as supererogatory acts). Praise is an often-used metric in psychology for assessing people's positive attitudes for moral actions [5–7,9–11]. Philosophers have referred to a praiseworthy action as one that is 'laudable' [12] and state that a person is praiseworthy when they perform a morally good action for morally worthy motives [13]. Other accounts of praise have emphasized the importance of the character of the person performing the moral action, highlighting how even the most moral act possible is not praiseworthy if performed for the wrong sorts of motivations [14,15]. We are restricting our discussion in this paper to moral praise, in order to distinguish it from the large literature on achievement praise (e.g., praise for academic or athletic performance). We are restricting our discussion to the judgment or appraisal of praise (and blame) and setting aside the instrumental role that praise (moral or achievement) may play in reinforcement learning. Finally, while we seek to highlight general distinctions between the psychological processes guiding blame and praise, there are notable exceptions to the patterns we identify and we note that both blame and praise are likely to operate in diverse ways across different content domains (Box 1).

Recent work has documented several asymmetries in the psychological processes guiding judgments of praise versus those of blame. For instance, people are differentially sensitive to outcome-magnitude: they are less sensitive to benefits of a positive act when making praise judgments than they are to the degree of harm caused by an act when attributing blame [7]. In addition, while individuals are demonstrably sensitive to criteria such as the intentionality and

Highlights

Recent psychological research suggests that the process for assigning moral praise is different than the process for assigning moral blame. Specifically, features of a moral act like causality, intentionality, and magnitude of outcome influence blame more than praise.

A functional perspective can help explain these differences while also offering a fruitful means for guiding future research into moral judgments. While blame is for determining who to avoid and signaling one's own moral character, praise is more for establishing cooperative alliances and relationships.

Praise is especially sensitive to factors that signal that a moral agent is a worthy cooperative partner and is guided by prosocial motivations.

Understanding the psychological process for moral praise offers insight into the development and evolution of cooperation.

[1]Department of Psychology, Cornell University, Ithaca, NY, USA
[2]Department of Psychology, Yale University, New Haven, CT, USA

*Correspondence:
raa255@cornell.edu (R.A. Anderson).

> **Box 1. Nuances in Attributions of Moral Praise and Blame**
>
> Although we have characterized praise and blame as tending to operate in certain ways for ease of understanding (e.g., blame as more susceptible to information regarding causality, intentionality, and magnitude than praise), we acknowledge that this is not always the case. The judgments we call 'moral praise' and 'moral blame' likely arise from diverse and complex cognitive mechanisms that depend on the particular actions and agents being evaluated and the cultural contexts in which they occur. For example, actions pertaining to the purity domain (i.e., actions concerned with moral, spiritual, and sexual cleanliness) appear to be distinguishable from other sorts of moral concerns and operate according to different cognitive processes. Blame for violations of moral purity are often insensitive regarding intentionality, such that impure acts are judged to be just as wrong when committed intentionally as unintentionally [72,119]. In addition, magnitude of outcomes appears to influence blame for harmful acts (e.g., causing more pain is worse than causing less pain) more than blame for impure acts (e.g., doing something sexually taboo or disgusting once is often equivalent to doing it many times [120]). Finally, moral actions made by members of a group described as 'impure' lead to amplification of both blame and praise [121]. Neither praise nor blame are the product of a single unified cognitive process; instead, both judgments are likely an amalgamation of several different, context-sensitive processes. It is important for future research to investigate both what distinguishes praise from blame, as well as to investigate the similarities and differences between different types of praise (see Outstanding Questions).

controllability of an action when determining whether to assign blame, these factors appear to play a smaller role in attributions of praise [5,6]. Importantly, however, it is not simply that individuals are less sensitive to contextual information in general when making judgments of praise [16–18]. Rather, people do appear to be especially sensitive to information regarding the motives driving the good deeds of others; even exceptionally generous acts are not seen as praiseworthy if they are viewed as motivated by self-interest [19]. Taken together, we believe that this work suggests that (more so than judgments of blame), praise judgments operate like an answer to the question 'is this a good person?' rather than 'is this a good action?'.

Here, we synthesize a growing literature focused on the evaluation of moral character [16,17] to sketch the foundations of a theory of moral praise. Recent research highlights the importance of character in our social evaluations, suggesting that we make judgments not just of particular acts but of the people who commit those acts [18]. Specifically, observers attend to information that reveals not just what a person did but why they did it in order to understand their underlying moral motivations and attitudes [5,10,19–21]. Such person-centered moral evaluations have important consequences: for example, they more strongly predict liking and respect for an individual than judgments of competence and sociability [22]. We propose that person-centered moral evaluations play an especially important role in moral praise. We suggest that this is because praise, relative to blame, is more geared towards relationship-building, and that people should be especially motivated to build relationships with those perceived to have good moral character. This perspective on the function of praise helps make sense of observed asymmetries between praise and blame.

Theories of Moral Responsibility

Traditionally, discussions of blame and praise have occurred in the context of determining the conditions under which people are held morally responsible for their actions: as having intentionally caused those actions and can thus be held accountable for the outcomes. While a great deal of research on moral judgment has focused on how we determine whether a particular action is right or wrong, when an individual engages in a morally relevant action, an additional moral judgment is required to determine whether he or she should be held morally responsible. How do people make these judgments? A number of highly influential theoretical accounts have been proposed [3,4]. These theories have been important in delineating the psychological processes involved in assigning blame, highlighting the importance of causality [1,2,23–25], intentionality [1,2,26,27], and the magnitude/severity of the outcome [1,2]. If an individual engages in an immoral act, for instance, these theories hold that we ask ourselves a series of questions before arriving at a judgment of responsibility and blame: did the person cause the action (if the agent is not causally linked to the outcome, there should be no blame [1,28,29])? Did they intend

the action (agents are more likely to be deemed blameworthy when they acted intentionally [1,2,30–32])? If so, did they have a good reason for acting? And how severe were the consequences of their action [1,33,34]? Notably, these theories have been almost exclusively concerned with judgments of responsibility for immoral acts and the resulting attributions of blame for those acts. Judgments of praise for positive acts have been assumed, either tacitly or explicitly, to result from the same set of attributional processes [2–4]. However, as we discuss in greater detail later, considerations of causality, intentionality, and consequences seem to matter less for praise than they do for blame.

The Functions of Blame and Praise

We propose that considering the ultimate functions and the potential consequences (both costs and benefits) of blame and praise can help shed light on the nature of moral praise and can help explain how and why judgments of blame and praise differ in very specific ways. Here, we adopt the view (proposed by many researchers) that moral judgments serve to facilitate social regulation and sustain social relationships, ensuring cooperation among group members [35–39]. Viewed through this lens, blame and praise can be understood as judgments that play related, but distinct, roles in social regulation. For instance, the expression of both praise and blame can help shape an agent's future moral behavior [3,35,36], can help identify which individuals are worthy social partners [2,3,28,40], and can publicly signal the moral values of the person making the judgment [41,42]. However, we propose that, compared with blame, praise is relatively more directed towards building, establishing, and maintaining social relationships and affiliative alliances. Because being praised signals a target's social value, praise works to improve interpersonal commitment [43], group commitment [44], and the social reputation of the recipient [45]. These relationship-building benefits can also result in wider application of praise than blame [46].

Praise and blame also differ in another important way: there are different costs associated with making an incorrect judgment. The wrongful attribution of blame can have very serious consequences, from social exclusion to severe physical punishment for the person who is wrongly blamed. Blame that is considered unjustified or miscalibrated can result in resentment [47] and decreased group commitment [48] from the person being blamed and psychological harm (from being unjustly punished) [49]. Even blame that is viewed as deserved by third parties may feel undeserved by the recipient, given that people often judge their own misdeeds as relatively less blameworthy than the misdeeds of others [50,51]. Additionally, wrongfully blaming a person may result in a damaged or lost relationship, or in possible retaliation from the accused [52,53]. Finally, even if an agent is deserving of blame, the person expressing a judgment can open herself to charges of hypocrisy if she were to engage in that same behavior [42]. Relative to blame, praise is comparatively costless: people typically enjoy being praised and the consequences for wrongfully praising an agent are relatively minimal for both the judge and the recipient of praise. Accordingly, blame often requires justification that praise simply does not [54]. It therefore becomes important to calibrate judgments of blame to inputs like the causality and intentionality of the agent, as well as to the magnitude of consequences of the action. In short, getting blame right matters more.

Despite the importance of getting the details about an action right when assessing blame, a growing body of work suggests that evaluations of the immoral acts of an agent are not only sensitive to these local features of an action (such as intentionality, control, and causality), but also to the character of an agent [1–5,7,16–18]. We propose that this is even more true for judgments of praise: because the cost of praise is less severe than the cost of blame, praise is even less sensitive to the fine-grained analyses of action that guide judgments of blame [8]. This does not mean that praise is entirely insensitive to these local features (it is rare that one would praise an entirely accidental positive action [25]), but only that it will be relatively less sensitive to information about these features. We believe that this

difference is a key to understanding the empirical findings regarding praise that have emerged in recent years.

Asymmetries in the Attribution of Praise and Blame

Judgments of blame and praise, we have argued, differ in important ways: while both are sensitive to the character of the agent, blame requires a degree of sensitivity to the specific features of an action that praise does not. On this view, compared with judgments of blame, judgments of praise ought to demonstrate a relative insensitivity to features such as the causal influence over the act, the intentionality of an action, or the magnitude of any given outcome (for exceptions, see Box 1). Within the past few years, a number of asymmetries between blame and praise have been documented that are consistent with these predictions. While, as we have noted, praise has received far less attention in the literature than blame, our review of the literature revealed at least 22 published papers comparing praise and blame (or evaluations of moral and immoral events more broadly) [5–9,25,55–70]. Of these, 20 of the papers documented a number of asymmetries consistent with our theoretical framework (Table 1). It is to these specific findings that we now turn.

Table 1. Published Papers Comparing Praise and Blame

Citation	Number of studies	Investigation of causality, intentionality, or magnitude of consequences?	Number of studies documenting a positive–negative asymmetry in judgments	Content domains of the moral acts being judged	Refs
Pizarro et al. (2003)	3	Intentionality	3	Helping and harming	[5]
Ohtsubo (2007)	2	Intentionality	2	Helping and harming	[6]
Siegel et al. (2017)	2	Causality, magnitude	2	Self's profit and pain versus other's profit and pain	[7]
Guglielmo and Malle (2019)	4	Intentionality	4	Helping and harming, purity	[8]
Wiltermuth et al. (2010)	3	None (condemnation and praise are orthogonal judgments)	3	Helping and harming, fairness, honesty	[9]
Pizarro et al. (2003)	4	Causality	0	Helping and harming	[25]
Bohner et al. (1988)	1	Causality	1	Success versus failure on task	[55]
Roese and Olson (1997)	Review article	Causality	Argues for asymmetry in causality based on valence	Success versus failure on task, helping and harming	[56]
Bostyn and Roets (2016)	3	Causality	2	Helping and harming	[57]
Newman et al. (2015)	5	Intentionality	5	Social norms, self-control, helping and harming	[58]
Knobe (2003)	2	Intentionality	2	Helping and harming	[59]
Knobe (2003)	4	Intentionality	4	Helping and harming	[60]
Leslie et al. (2006)	2	Intentionality	2	Helping and harming	[61]
Ngo et al. (2015)	3	Intentionality	3	Helping and harming	[62]
Klein and Epley (2014)	8	Magnitude	5	Economic decisions	[63]
Gneezy and Epley (2014)	6	Magnitude	6	Honesty, keeping of promises	[64]
Klein and O'Brien (2016)	5	Causality, magnitude	5	Nice versus mean behavior, selfish versus selfless behavior	[65]
Goodwin and Darley (2012)	2	None (perceived objectivity)	1	Helping and harming, respect, honesty, donation	[66]
Critcher et al. (2013)	2	Intentionality	1	Helping and harming, purity	[67]
Bigman and Tamir (2016)	7	None (effort)	0	Helping and harming	[68]
Cushman et al. (2009)	1	Intentionality	1	Self versus other profit	[69]
Monroe et al. (2018)	5	Magnitude	4	Social norms, economic decisions	[70]

Causality

Compared with judgments of moral blame, judgments of moral praise appear less attuned to the degree of causality that an agent had over an outcome (i.e., an evaluation that were it not for the actions of this agent, then these harmful consequences would not have occurred) [1,28,29]. More generally, negative events generate greater consideration of the causal mechanisms that led to the event [55] and give rise to more counter-factual thinking (i.e., 'what if' statements) than positive events [56], suggesting that people are more motivated to understand the causes of immoral acts than of moral acts. Moreover, prompting participants to consider causality appears to lead to increases in praise for the agent, but not to changes in blame for the agent, suggesting that moral events fail to spontaneously trigger causal attribution to the same degree as immoral events [55].

As further evidence of the difference in sensitivity to causality when making judgments of blame versus praise, there is evidence that the so-called 'omission bias' is less likely to occur for positive moral actions than it is for negative moral actions: individuals are more likely to reduce blame for omissions compared with actions than they are to reduce praise. For example, in one study participants assigned greater blame to an individual who moved out of the way of a cart that subsequently hit other people, compared with an individual who failed to jump in front of a cart in order to prevent it from hitting others. However, participants assigned similar levels of praise for an individual who jumped in front of a cart to prevent it from hitting others and an individual who did not move out of the way (staying in the cart's path) in order to prevent it from hitting others [57]. These findings are consistent with the approach we have proposed: blame requires greater justification than praise due to the potential costs of getting it wrong [54].

Intentionality

A consistent finding (with a few exceptions [71,72]) within the field of moral psychology is that agents who commit immoral acts are considered more blameworthy when they acted intentionally (i.e., with accurate beliefs about and desires for the action and its consequences) than when they acted unintentionally [1–4,23,27,73]. However, research has also consistently shown that observers treat the intentionality of agents differently when evaluating moral acts versus immoral acts. For instance, intentionality information appears to matter more for blame than for praise. Observers discount blame for impulsive (and therefore less intentional) immoral acts relative to deliberative immoral acts but fail to discount praise for impulsive moral acts relative to deliberative moral acts because of an assumption regarding what the agent would 'truly' want to happen [5,58]. Similarly, providing individuals with explicit information about intentionality intensifies the blame they assign to a harmful agent more strongly than such information intensifies the praise they assign to a helpful agent [6,8].

The motivations of agents who commit positive moral acts are also judged as more ambiguous than the motivations of agents who commit immoral acts [74,75]. For instance, people are more likely to attribute immoral motives to an agent who acts morally than they are to attribute moral motives to an agent who acts immorally [19,76]. This suggests that observers are particularly interested in understanding whether an agent is acting morally for genuinely prosocial reasons and not just trying to appear moral in order to reap reputational benefits, a finding that is consistent with viewing praise as important to judgments of character. Given the motivational ambiguity surrounding an agent's positive moral acts, observers are especially scrutinizing of a prosocial agent's motives.

The well-documented 'side-effect effect' in judgments of intentionality also highlights this difference between moral and immoral acts. Across a number of studies, researchers have documented that

when an action has an explicitly unintentional negative side-effect, individuals nonetheless judge that the side-effect was intentional; however, when an action has an explicitly unintentional positive side-effect, individuals are much less likely to judge that the side-effect was intentional [59–62]. These findings provide additional evidence that the relationship between the intentionality of an act and the harm caused by that act is different than the relationship between the intentionality of an act and the benefit caused by that act.

Like information about causality, information about intentionality is not irrelevant for judgments of praise. Rather, it appears as if information about the intentionality of a single moral act is simply less important for judgments of praise than is information about an agent's enduring moral motivations. This is consistent with the view that the function of praise is establishing and maintaining social bonds. While information about the intentionality of an action often serves as a useful signal of an agent's underlying motivations, this is more true for negative than for positive actions. Because people assume that others generally want good things to happen [5,58], clear deviations from this norm (i.e., when an individual intends for a bad thing to happen) are particularly diagnostic of an individual's motivation [75]. Finally, as we argued earlier, careful considerations of intentionality may be more likely to occur for negative acts because of concerns over incorrectly blaming an agent who does not deserve it. Such concerns may be less likely to arise when making judgments of praise.

Magnitude of Outcomes

A third asymmetry between blame and praise has been documented regarding the magnitude of consequences for an act. While acts with more harmful outcomes are given greater blame than acts with less harmful outcomes [1,77,78], the magnitude of outcomes matters less for the evaluation of moral acts than immoral acts [70]. While individuals do use information about outcome magnitude when they are simultaneously evaluating positive acts that differ in magnitude, they fail to do so when considering individual acts in isolation [63]. For example, people view a player as increasingly moral as they move from making completely selfish allocations in an economic game to making more equitable allocations, but they do not continue to view them as increasingly more moral as they move equitable allocations to completely selfless ones [63]. Similarly, people judge a broken promise as worse than a kept promise but judge an exceeded promise as no better than a kept promise [64]. People also require less evidence to diagnose moral change for the worse than to diagnose moral change for the better [65], suggesting that observers are more sensitive to the gradations of outcomes of immoral acts than for moral acts. Finally, observers seem to weight outcomes more strongly when judging agents who commit more negative than positive acts [7,8]. In short, when evaluating the actions of a moral agent, people seem to care more about the act of doing good than how much good was done [10].

One potential explanation for this asymmetry in sensitivity to outcome magnitude is that people broadly judge bad stimuli as stronger than good stimuli and negative information often has more impact than equally positive information [79–82]. For example, people are generally loss averse, and see losing $5 as hurting more than gaining $5 feels good [83]. People also give much greater weight to negative events than positive events in their evaluations and attitudes [84]. Furthermore, moral wrongs are seen as more objectively wrong than moral goods are seen as objectively right [66], providing additional evidence that the valence of a behavior influences judgments of its strength. This general 'negativity bias' appears to emerge early in life; for instance, both children and adults have greater memory for threatening versus nonthreatening others [85]. An important question for future research is whether the observed asymmetries between blame and praise in sensitivity to outcome magnitude can be entirely explained by loss aversion and negativity bias, or whether such asymmetries additionally arise due to the fact that attribution errors are more costly for blame than praise.

What Matters for Praise

We propose that when determining whether to morally praise another person, observers are trying to determine whether that person is a fundamentally trustworthy individual, one who can be counted on to be a cooperative partner in the future. Praise then acts to reinforce that moral behavior in the recipient while also strengthening the relationship between the giver and the recipient of the praise. When faced with positive acts, observers appear to be both diagnosing a person's moral character and generating a prediction about their future prosocial behavior. Praise should therefore be sensitive to features of acts and persons indicating that a person can be reliably trusted to act prosocially. Praise is a judgment about an agent's 'moral cognitive machinery' [16]; an evaluation that the agent will, if placed in a situation where morality is relevant, make the appropriate (i.e., normatively prescribed) moral judgments and decisions. While information about causality [5], intentionality [86], and outcome magnitude [63] of an act are relevant in evaluating the praiseworthiness of an act, those factors are only important to the extent they inform judgments of character. Other features of an act or an agent that serve as a clearer signal about the agent's moral character should, on our account, be especially important for the attribution of praise.

For example, the emotions an agent displays when making a decision are used as cues to that agent's motivations and moral character for both helpful and harmful actions: moral actions performed with a happy expression, relative to a less happy expression, communicate the agent's moral commitments [87]. Observers infer character from the emotions displayed by an agent because affective displays are assumed to communicate the agent's motives and intentions [88,89]; if an agent smiles while behaving morally, observers assume that the agent approves of what is happening, more than if the agent is frowning. Furthermore, observers change how they interpret an agent's helpful behavior based on the agent's degree of willingness to help and the apparent reasons for helping: if an agent appears to only grudgingly help [90] or to help solely based on cost-benefit analysis [91], judges award less praise to the agent. People see emotional expressions and 'warm-glow' feelings after behaving prosocially as signals of an agent's moral character [20,92] and, correspondingly, 'warm glow' feelings are relied upon more strongly when assigning praise than blame [10]. However, the reputational and social benefits of positive expressions over neutral expressions diminish with repeated behavior [87], suggesting that displayed emotions are treated as signaling an agent's prosocial commitment, but that actual patterns of behavior are a clearer signal of such commitment.

Similarly, individuals use the speed of a moral decision as diagnostic of moral character and the praiseworthiness of the decision [67]: observers judge hesitant decisions as less diagnostic and reflecting more ambiguous motivations relative to quick decisions when praising moral acts or when blaming immoral acts. Similarly, people believe that spontaneous (presumably quickly occurring) decisions reveal more about an agent's true desires than deliberative (presumably slowly occurring) decisions [93]. In economic games, people judge another's quick decisiveness and lack of calculation when making cooperation decisions as signaling trustworthiness [94–96]. To observers, quick moral decisions signal that an agent is certain about the moral decision, seemingly uninfluenced by situational factors, and can be reliably trusted on to cooperate and behave morally in the future.

In assigning praise, observers evaluate whether the agent appears to possess the right sort of motivations for acting morally and to predict that they will cooperate and behave trustworthily in the future. Emotional expressions and decision speed are proxies for such motivations, but more direct evidence of an individual's motivations for a moral decision can likewise influence praise. Although motives and reasons influence blame and can act as mitigating circumstances [2], research suggests that praise is more sensitive to mixed motivations than blame. For

instance, information that an agent is behaving morally for immoral (or at least amoral) reasons reduces attributions of praise. People value prosocial motivations and the absence of motivational temptation in an agent when praising that agent [97,98]. In some cases, observers attribute blame to individuals for just thinking negative thoughts [99,100]. Observers also assign less praise to an agent if the agent receives personal benefit for an altruistic behavior [101], even compared with not having performed the moral behavior at all [102], unless the personal benefit is emotional [97]. In addition, observers assume selfish motivations on the part of a moral agent and assign less praise if the agent brags about public prosocial behavior [103].

The amount of effort exerted in performing either a moral act or an immoral act can similarly serve as a signal about an agent's underlying character [68], because effort is assumed to reflect the relative importance of the goal to the agent [104–108]. Individuals are likely to infer goals when observing effortful behavior [109] and greater effort exerted by an agent leads to stronger inferences about the goals of that agent [110,111]. This tendency to infer goals from effort can be observed in young children [112]. Moral praise is also more likely to be given to agents whose prosocial behavior was effortful and lacked ease [113] (although this sensitivity to effort only emerges later in development [114]) and for acts that are judged as extraordinary (compared with simply ordinary [115]) and exceeded their obligations and duties [116]. Similarly, acts that are costly, or that require sacrifice for others, are judged as more morally heroic and praiseworthy [117,118].

Concluding Remarks

Moral praise, we have argued, is a psychological response that, like other forms of moral judgment, serves a particular functional role in establishing social bonds, encouraging cooperative alliances, and promoting good behavior. Through this lens, seemingly perplexing asymmetries between judgments of blame for immoral acts and judgments of praise for moral acts can be understood as consistent with the relative roles, and associated costs, played by these two kinds of moral judgments. While both blame and praise judgments require that an agent played some causal and intentional role in the act being judged, praise appears to be less sensitive to these features and more sensitive to more general features about an individual's stable, underlying character traits. In other words, we believe that the growth of studies on moral praise in the past few years demonstrate that, when deciding whether or not doling out praise is justified, individuals seem to care less on how the action was performed and far more about what kind of person performed the action. We suggest that future research on moral attribution should seek to complement the rich literature examining moral blame by examining potentially unique processes engaged in moral praise, guided by an understanding of their differing costs and benefits, as well as their potentially distinct functional roles in social life (see Outstanding Questions).

References

1. Cushman, F. (2008) Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition* 108, 353–380
2. Malle, B.F. *et al.* (2014) A theory of blame. *Psychol. Inq.* 25, 147–186
3. Shaver, K.G. (1985) *The Attribution of Blame: Causality, Responsibility, and Blameworthiness*, Springer Verlag
4. Weiner, B. (1995) *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*, Guilford
5. Pizarro, D. *et al.* (2003) Asymmetry in judgments of moral blame and praise the role of perceived metadesires. *Psychol. Sci.* 14, 267–272
6. Ohtsubo, Y. (2007) Perceived intentionality intensifies blameworthiness of negative behaviors: blame-praise asymmetry in intensification effect. *Jpn. Psychol. Res.* 49, 110
7. Siegel, J.Z. *et al.* (2017) Inferences about moral character moderate the impact of consequences on blame and praise. *Cognition* 167, 201–211
8. Guglielmo, S. and Malle, B.F. (2019) Asymmetric morality: Blame is more differentiated and more extreme than praise. *PLoS One* 14, e0213544
9. Wiltermuth, S.S. *et al.* (2010) The orthogonality of praise and condemnation in moral judgment. *Soc. Psychol. Pers. Sci.* 1, 302–310
10. Yudkin, D.A. *et al.* (2019) Actions speak louder than outcomes in judgments of prosocial behavior. *Emotion* 19, 1138–1147
11. Nelkin, D.K. (2016) Difficulty and degrees of moral praiseworthiness and blameworthiness. *Noûs* 50, 356–378
12. Smith, H.M. (1991) Varieties of moral worth and moral credit. *Ethics* 101, 279–303
13. Arpaly, N. and Schroeder, T. (1999) Praise, blame, and the whole self. *Philos. Stud.* 93, 161–188
14. Archer, A. (2016) Are acts of supererogation always praiseworthy? *Theoria* 82, 238–255
15. McNamara, P. (2011) Supererogation, inside and out: toward an adequate scheme for common-sense morality. In *Oxford Studies*

in Normative Ethics (Vol. 1) (Timmons, M., ed.), pp. 202–235, Oxford University Press

16. Helzer, E.G. and Critcher, C.R. (2018) What do we evaluate when we evaluate moral character? In *Atlas of Moral Psychology* (Gray, K. and Graham, J., eds), pp. 99–107, Guilford Press

17. Uhlmann, E.L. *et al.* (2015) A person-centered approach to moral judgment. *Perspect. Psychol. Sci.* 10, 72–81

18. Goodwin, G.P. *et al.* (2014) Moral character predominates in person perception and evaluation. *J. Pers. Soc. Psychol.* 106, 148–168

19. Critcher, C.R. and Dunning, D. (2011) No good deed goes unquestioned: cynical reconstruals maintain belief in the power of self-interest. *J. Exp. Soc. Psychol.* 47, 1207–1213

20. Barasch, A. *et al.* (2014) Selfish or selfless? On the signal value of emotion in altruistic behavior. *J. Pers. Soc. Psychol.* 107, 393–413

21. Uhlmann, E.L. *et al.* (2014) When actions speak volumes: the role of inferences about moral character in outrage over racial bigotry. *Eur. J. Soc. Psychol.* 44, 23–29

22. Hartley, A.G. *et al.* (2016) Morality's centrality to liking, respecting, and understanding others. *Soc. Psychol. Pers. Sci.* 7, 648–657

23. Shaver, K.G. and Drown, D. (1986) On causality, responsibility, and self-blame: a theoretical note. *J. Pers. Soc. Psychol.* 50, 697–702

24. Fincham, F.D. and Roberts, C. (1985) Intervening causation and the mitigation of responsibility for harm doing. *J. Exp. Soc. Psychol.* 21, 178–194

25. Pizarro, D.A. *et al.* (2003) Causal deviance and the attribution of moral responsibility. *J. Exp. Soc. Psychol.* 39, 653–660

26. Darley, J.M. *et al.* (1978) Intentions and their contexts in the moral judgments of children and adults. *Child Dev.* 49, 66–74

27. Darley, J.M. and Shultz, T.R. (1990) Moral rules – their content and acquisition. *Annu. Rev. Psychol.* 41, 525–556

28. Sloman, S.A. *et al.* (2009) Causal models: the representational infrastructure for moral judgment. In *Moral Judgment and Decision Making* (Bartels, D. *et al.*, eds), pp. 1–26, Academic Press

29. Shultz, T.R. *et al.* (1986) Assignment of moral responsibility and punishment. *Child Dev.* 57, 177–184

30. Malle, B.F. and Knobe, J. (1997) Which behaviors do people explain? A basic actor-observer asymmetry. *J. Pers. Soc. Psychol.* 72, 288–304

31. Gray, K. and Wegner, D.M. (2008) The sting of intentional pain. *Psychol. Sci.* 19, 1260–1262

32. Karniol, R. (1978) Children's use of intention cues in evaluating behavior. *Psychol. Bull.* 85, 76–85

33. Miller, R.M. *et al.* (2014) Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion* 14, 573–587

34. Cushman, F. (2013) Action, outcome and value: a dual-system framework for morality. *Personal. Soc. Psychol. Rev.* 17, 273–292

35. Curry, O.S. (2016) Morality as cooperation: a problem-centred approach. In *The Evolution of Morality* (Shackelford, T.K. and Hansen, R.D., eds), pp. 27–51, Springer

36. Flack, J.C. and De Waal, F.B. (2000) 'Any animal whatever': Darwinian building blocks of morality in monkeys and apes. *J. Consciousness Stud.* 7, 1–29

37. Haidt, J. (2008) Morality. *Perspect. Psychol. Sci.* 3, 65–72

38. Joyce, R. (2006) *The Evolution of Morality*, MIT Press

39. Rai, T.S. and Fiske, A.P. (2011) Moral psychology is relationship regulation: moral motives for unity, hierarchy, equality, and proportionality. *Psychol. Rev.* 118, 55–75

40. Pizarro, D.A. and Tannenbaum, D. (2011) Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In *The Social Psychology of Morality: Exploring the Causes of Good and Evil* (Shaver, P. and Mikulincer, M., eds), pp. 91, APA Books

41. Jordan, J. *et al.* (2016) Third-party punishment as a costly signal of trustworthiness. *Nature* 530, 473–476

42. Jordan, J.J. *et al.* (2017) Why do we hate hypocrites? Evidence for a theory of false signaling. *Psychol. Sci.* 28, 356–368

43. Algoe, S.B. *et al.* (2016) Putting the "you" in "thank you": examining other-praising behavior as the active relational ingredient in expressed gratitude. *Soc. Psychol. Pers. Sci.* 7, 658–666

44. Eisenberger, R. *et al.* (1986) Perceived organizational support. *J. Appl. Psych.* 71, 500–507

45. Henrich, J. and Gil-White, F.J. (2001) The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* 22, 165–196

46. Schein, C. *et al.* (2020) Praise-many, blame-fewer: a common (and successful) strategy for attributing responsibility in groups. *J. Exp. Psych. Gen.* 149, 855–869

47. Aquino, K. *et al.* (2001) How employees respond to personal offense: the effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace. *J. Appl. Psych.* 86, 52–59

48. Podsakoff, P.M. *et al.* (2006) Relationships between leader reward and punishment behavior and subordinate attitudes, perceptions, and behaviors: a meta-analytic review of existing and new research. *Organ. Behav. Hum. Dec.* 99, 113–142

49. Grounds, A. (2004) Psychological consequences of wrongful conviction and imprisonment. *Can. J. Criminol. Crim.* 46, 165–182

50. Elshout, M. *et al.* (2017) Your act is worse than mine: perception bias in revenge situations. *Aggressive Behav.* 43, 553–557

51. Stillwell, A.M. *et al.* (2008) We're all victims here: toward a psychology of revenge. *Basic Appl. Soc. Psych.* 30, 253–263

52. Dreber, A. *et al.* (2008) Winners don't punish. *Nature* 452, 348–351

53. McCullough, M.E. *et al.* (2013) Cognitive systems for revenge and forgiveness. *Behav. Brain Sci.* 36, 1–15

54. Voiklis, J. and Malle, B.F. (2017) Moral cognition and its basis in social cognition and social regulation. In *Atlas of Moral Psychology* (Gray, K. and Graham, J., eds), pp. 108–120, Guilford Press

55. Bohner, G. *et al.* (1988) What triggers causal attributions? The impact of valence and subjective probability. *Eur. J. Soc. Psychol.* 18, 335–345

56. Roese, N.J. and Olson, J.M. (1997) Counterfactual thinking: the intersection of affect and function. *Adv. Exp. Soc. Psychol.* 29, 1–59

57. Bostyn, D.H. and Roets, A. (2016) The morality of action: the asymmetry between judgments of praise and blame in the action–omission effect. *J. Exp. Soc. Psychol.* 63, 19–25

58. Newman, G.E. *et al.* (2015) Beliefs about the true self explain asymmetries based on moral judgment. *Cognitive Sci.* 39, 96–125

59. Knobe, J. (2003) Intentional action and side effects in ordinary language. *Analysis* 63, 190–194

60. Knobe, J. (2003) Intentional action in folk psychology: an experimental investigation. *Philos. Psychol.* 16, 309–324

61. Leslie, A.M. *et al.* (2006) Acting intentionally and the side-effect effect: theory of mind and moral judgment. *Psychol. Sci.* 17, 421–427

62. Ngo, L. *et al.* (2015) Two distinct moral mechanisms for ascribing and denying intentionality. *Sci. Rep.* 5, 17390

63. Klein, N. and Epley, N. (2014) The topography of generosity: asymmetric evaluations of prosocial actions. *J. Exp. Psych. Gen.* 143, 2366–2379

64. Gneezy, A. and Epley, N. (2014) Worth keeping but not exceeding: asymmetric consequences of breaking versus exceeding promises. *Soc. Psychol. Pers. Sci.* 5, 796–804

65. Klein, N. and O'Brien, E. (2016) The tipping point of moral change: when do good and bad acts make good and bad actors? *Soc. Cognition* 34, 149–166

66. Goodwin, G.P. and Darley, J.M. (2012) Why are some moral beliefs perceived to be more objective than others? *J. Exp. Soc. Psychol.* 48, 250–256

67. Critcher, C.R. *et al.* (2013) How quick decisions illuminate moral character. *Soc. Psychol. Pers. Sci.* 4, 308–315

68. Bigman, Y.E. and Tamir, M. (2016) The road to heaven is paved with effort: perceived effort amplifies moral judgment. *J. Exp. Psychol. Gen.* 145, 1654–1669

69. Cushman, F. *et al.* (2009) Accidental outcomes guide punishments in a "trembling hand" game. *PLoS One* 4, e6699

70. Monroe, A.E. *et al.* (2018) It's not what you do, but what everyone else does: on the role of descriptive norms and subjectivism in moral judgment. *J. Exp. Soc. Psychol.* 77, 1–10

71. Young, L. *et al.* (2010) Investigating the neural and cognitive basis of moral luck: it's not what you do but what you know. *Rev. Philos. Psychol.* 1, 333–349

72. Young, L. and Saxe, R. (2011) When ignorance is no excuse: Different roles for intent across moral domains. *Cognition* 120, 202–214

73. Cushman, F. *et al.* (2013) The development of intent-based moral judgment. *Cognition* 127, 6–21

74. Reeder, G.D. and Brewer, M.B. (1979) A schematic model of dispositional attribution in interpersonal perception. *Psychol. Rev.* 86, 61–79

75. Skowronski, J.J. and Carlston, D.E. (1987) Social judgment and social memory: the role of cue diagnosticity in negativity, positivity, and extremity biases. *J. Pers. Soc. Psychol.* 52, 689–699

76. Miller, D. (1999) The norm of self-interest. *Am. Psychol.* 54, 1053–1060

77. Kahneman, D. *et al.* (1998) Shared outrage and erratic awards: the psychology of punitive damages. *J. Risk Uncertainty* 16, 49–86

78. Walster, E. (1966) Assignment of responsibility for an accident. *J. Pers. Soc. Psychol.* 3, 73–79

79. Baumeister, R.F. *et al.* (2001) Bad is stronger than good. *Rev. Gen. Psychol.* 5, 323–370

80. Fiske, S.T. (1980) Attention and weight in person perception: the impact of negative and extreme behavior. *J. Pers. Soc. Psychol.* 38, 889–906

81. Taylor, S.E. (1991) Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychol. Bull.* 110, 67–85

82. Wentura, D. *et al.* (2000) Automatic vigilance: the attention-grabbing power of approach-and avoidance-related social information. *J. Pers. Soc. Psychol.* 78, 1024–1037

83. Tversky, A. and Kahneman, D. (1991) Loss aversion in riskless choice: a reference-dependent model. *Q. J. Econ.* 106, 1039–1061

84. Rozin, P. and Royzman, E.B. (2001) Negativity bias, negativity dominance, and contagion. *Personal. Soc. Psychol. Rev.* 5, 296–320

85. Kinzler, K.D. and Shutts, K. (2008) Memory for "mean" over "nice": the influence of threat on children's face memory. *Cognition* 107, 775–783

86. Weiner, B. and Kukla, A. (1970) An attributional analysis of achievement motivation. *J. Pers. Soc. Psychol.* 55, 738–748

87. Ames, D.R. and 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. *Psychol. Sci.* 20, 586–593

88. Higgins, E.T. (1998) The aboutness principle: a pervasive influence on human inference. *Soc. Cognition* 16, 173–198

89. Wellman, H.M. *et al.* (2000) Young children's understanding of perception, desire, and emotion. *Child Dev.* 71, 895–912

90. Krull, D.S. *et al.* (2008) Smile when you say that: effects of willingness on dispositional inferences. *J. Exp. Soc. Psychol.* 44, 735–742

91. Ames, D.R. *et al.* (2004) It's the thought that counts: on perceiving how helpers decide to lend a hand. *Pers. Soc. Psychol. B* 30, 461–474

92. Levine, E.E. (2018) Signaling emotion and reason in cooperation. *J. Exp. Psychol. Gen.* 147, 702–719

93. Morewedge, C.K. *et al.* (2014) The (perceived) meaning of spontaneous thoughts. *J. Exp. Psychol. Gen.* 143, 1742–1754

94. Evans, A.M. and van de Calseyde, P.P. (2017) The effects of observed decision time on expectations of extremity and cooperation. *J. Exp. Soc. Psychol.* 68, 50–59

95. Hoffman, M. *et al.* (2015) Cooperate without looking: why we care what people think and not just what they do. *Proc. Natl. Acad. Sci. U. S. A.* 112, 1727–1732

96. Jordan, J.J. *et al.* (2016) Uncalculating cooperation is used to signal trustworthiness. *Proc. Natl. Acad. Sci. U. S. A.* 113, 8658–8663

97. Carlson, R.W. and Zaki, J. (2018) Good deeds gone bad: Lay theories of altruism and selfishness. *J. Exp. Soc. Psychol.* 75, 36–40

98. Berman, J.Z. and Small, D.A. (2018) Discipline and desire: on the relative importance of willpower and purity in signaling virtue. *J. Exp. Soc. Psychol.* 76, 220–230

99. Inbar, Y. *et al.* (2012) Benefiting from misfortune: when harmless actions are judged to be morally blameworthy. *Pers. Soc. Psychol. B* 38, 52–62

100. Tetlock, P.E. *et al.* (2000) The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. *J. Pers. Soc. Psychol.* 78, 853–870

101. Barasch, A. *et al.* (2016) When payment undermines the pitch: on the persuasiveness of pure motives in fund-raising. *Psychol. Sci.* 27, 1388–1397

102. Newman, G.E. and Cain, D.M. (2014) Tainted altruism: when doing some good is evaluated worse than doing no good at all. *Psychol. Sci.* 25, 648–655

103. Berman, J.Z. *et al.* (2015) The braggart's dilemma: on the social rewards and penalties of advertising prosocial behavior. *J. Marketing Res.* 52, 90–104

104. Heider, F. (1958) *The Psychology of Interpersonal Relations*, Wiley

105. Henrich, J. (2009) The evolution of costly displays, cooperation and religion. *Evol. Hum. Behav.* 30, 244–260

106. Austin, J.T. and Vancouver, J.B. (1996) Goal constructs in psychology: structure, process, and content. *Psychol. Bull.* 120, 338–375

107. Hollenbeck, J.R. *et al.* (1989) An empirical examination of the antecedents of commitment to difficult goals. *J. Appl. Psychol.* 74, 18–23

108. Oettingen, G. *et al.* (2001) Self-regulation of goal setting: turning free fantasies about the future into binding goals. *J. Pers. Soc. Psychol.* 80, 736–753

109. Hassin, R.R. *et al.* (2005) Automatic goal inferences. *J. Exp. Soc. Psychol.* 41, 129–140

110. Dik, G. and Aarts, H. (2007) Behavioral cues to others' motivation and goal pursuits: the perception of effort facilitates goal inference and contagion. *J. Exp. Soc. Psychol.* 43, 727–737

111. Dik, G. and Aarts, H. (2008) I want to know what you want: How effort perception facilitates the motivation to infer another's goal. *Soc. Cognition* 26, 737–754

112. Jara-Ettinger, J. *et al.* (2015) Not so innocent: toddlers' inferences about costs and culpability. *Psychol. Sci.* 26, 633–640

113. Janoff-Bulman, R. *et al.* (2009) Proscriptive versus prescriptive morality: two faces of moral regulation. *J. Pers. Soc. Psychol.* 96, 521–537

114. Starmans, C. and Bloom, P. (2016) When the spirit is willing, but the flesh is weak: developmental differences in judgments about inner moral conflict. *Psychol. Sci.* 27, 1498–1506

115. Futamura, I. (2018) Is extraordinary prosocial behavior more valuable than ordinary prosocial behavior? *PLoS One* 13, e0196340

116. McManus, R.M. *et al.* (2020) What we owe to the family: the impact of special obligations on moral judgment. *Psychol. Sci.* 31, 227–242

117. Kraft-Todd, G.T. and Rand, D.G. (2019) Rare and costly prosocial behaviors are perceived as heroic. *Front. Psychol.* 10, 234

118. Visserman, M.L. *et al.* (2018) It's the motive that counts: perceived sacrifice motives and gratitude in romantic relationships. *Emotion* 18, 625–637

119. Chakroff, A. *et al.* (2016) When minds matter for moral judgment: Intent information is neurally encoded for harmful but not impure acts. *Soc. Cogn. Affect. Neurosci.* 11, 476–484

120. Rottman, J. and Young, L. (2019) Specks of dirt and tons of pain: dosage distinguishes impurity from harm. *Psychol. Sci.* 30, 1151–1160

121. Masicampo, E.J. *et al.* (2014) Group-based discrimination in judgments of moral purity-related behaviors: experimental and archival evidence. *J. Exp. Psychol. Gen.* 143, 2135–2152