



# Outrage fatigue? Cognitive costs and decisions to blame

Veerpal Bamrah<sup>1</sup> · C. Daryl Cameron<sup>2,3</sup> · Michael Inzlicht<sup>4,5</sup>

Accepted: 26 November 2022 / Published online: 17 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Across nine studies ( $N = 1672$ ), we assessed the link between cognitive costs and the choice to express outrage by blaming. We developed the *Blame Selection Task*, a binary free-choice paradigm that examines individuals' propensity to blame transgressors (versus an alternative choice)—either before or after reading vignettes and viewing images of moral transgressions. We hypothesized that participants' choice to blame wrongdoers would negatively relate to how cognitively inefficacious, effortful, and aversive blaming feels (compared to the alternative choice). With vignettes, participants approached blaming and reported that blaming felt more efficacious. With images, participants avoided blaming and reported that blaming felt more inefficacious, effortful, and aversive. Blame choice was greater for vignette-based transgressions than image-based transgressions. Blame choice was positively related to moral personality constructs, blame-related norms, and perceived efficacy of blaming, and inversely related to perceived effort and aversion of blaming. The BST is a valid behavioral index of blame propensity, and choosing to blame is linked to its cognitive costs.

**Keywords** Moral outrage · Blame · Choice · Effort · Efficacy

## Introduction

Outrage often seems easy. In 2013, when Justine Sacco, a senior director of corporate communications at holding company InterActiveCorp, tweeted an acerbic joke about contracting a sexually transmitted disease in Africa, she was swiftly labeled as racist and verbally attacked by thousands of outraged strangers, making her (for a while) the No. 1 worldwide trend on Twitter (Ronson, 2015). In 2017, when over 100 travellers were detained at the United States border by an executive order blocking civilians from seven

Muslim-majority countries, tens of thousands of American citizens gathered nationwide to voice their detestation towards former President Donald Trump and towards the policy (Diamond, 2017). In 2019, when Alabama citizen Marshaé Jones was shot in the stomach during a dispute with another woman, it was Jones who was charged with manslaughter—a decision that prompted both national support from pro-life groups and law enforcement officials, who blamed Jones for initiating and continuing the fight that resulted in the death of her 5-month old fetus, and national outcry from pro-choice groups, who blamed Alabama's racialized criminalization of women's pregnancies for Jones' sudden indictment (Silverman & Ellis, 2019).

All of these cases are examples of *moral outrage*, which are feelings of anger, contempt, and disgust directed towards a third-party for violating some moral standard of fairness, justice, or care (Haidt, 2003; Salerno & Peter-Hagene, 2013; Spring et al., 2018). As the examples above suggest, violations of fairness, justice, or care not only lead individuals to feel these other-condemning emotions, but to express them in different ways (Bastian et al., 2013). For example, spectators may express feelings of moral outrage by shaming or dehumanizing the individual(s) who they believe has violated a moral norm(s) (Crockett, 2017; Haslam, 2006; Viki et al., 2012). Alternatively, onlookers may call for retributive

✉ Veerpal Bamrah  
bamrahv@yorku.ca

<sup>1</sup> Department of Psychology, Faculty of Health, York University, Behavioural Sciences Building, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

<sup>2</sup> Department of Psychology, The Pennsylvania State University, University Park, PA 16801, USA

<sup>3</sup> Rock Ethics Institute, The Pennsylvania State University, University Park, PA 16801, USA

<sup>4</sup> Department of Psychology, University of Toronto, 1265 Military Trail, Scarborough, ON M1C 1A4, Canada

<sup>5</sup> Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, Canada

forms of punishment that are not only proportionate to the seriousness of moral wrong committed, but also to the intensity of their hostile affect towards the perpetrator (Carlsmith et al., 2002; Salerno & Peter-Hagene, 2013).

Another way of demonstrating outrage is by blaming, that is, assigning moral responsibility and culpability to the individual(s) for a fault or wrong (Alicke, 2000; Bastian et al., 2011). Blaming is a particularly important form of outrage expression, as some accounts of blame (e.g., Guglielmo et al., 2009; Malle et al., 2014; Monroe & Malle, 2017) suggest that onlookers sequentially process ‘agent’ causation, intent, reasons, desires, capacity/foresight, and blameworthiness for a norm-violating event *before* shaming, dehumanizing, and punishing the agent(s). That is, the assignment and degree of blame is said to precede (and thus, justify) other forms of outrage expression. It follows, then, that understanding the causes and consequences of expressions of moral outrage should begin with exploring blame.

Now more than ever, individuals are readily able to express their outrage. Online, users can express their outrage directly to the transgressor or to a broader audience with just a few keystrokes, anonymously or openly; and with even less work, they can simply repost or react to others’ comments, fueling virality (Brady et al., 2017; Brady et al., 2020a, b; Crockett, 2017). Yet, regardless of what is expressed (blaming, shaming, dehumanizing, or punishing sentiments), moral outrage has often been cast as exhausting: With the deluge of offensive news stories found online, as opposed to in-person or through traditional forms of media (Crockett, 2017; Hofmann et al., 2014), some have suggested that *outrage fatigue* can prevent individuals from being morally engaged and from taking pro-social or collective actions (Barberá et al., 2015; Brady et al., 2017; Crockett, 2017; Xiao & Houser, 2005). These conflicting intuitions suggest there is a need to more completely understand the predictors and consequences of feeling and expressing moral outrage. We developed a task that uses behaviorally revealed preferences (Kool et al., 2010) to empirically explore the propensity to blame—that is, whether individuals will generate outrage and blame others (i.e., assign responsibility for a fault or wrong) when given the choice to blame against an alternative act. Our behavioral approach compliments recent individual differences approaches to understanding outrage/blame propensity (see Gill & Cerce, 2021). Additionally, and crucially, in order to understand whether expressing outrage is *cognitively* fatiguing, we empirically examined the felt cognitive costs of blaming and how, if at all, such costs relate to individuals’ choices to blame others.

### **Motivations for blame (and other expressions of moral outrage)**

Philosophers, sociologists, legal theorists, and experimental psychologists are largely in agreement as to what comprises

a blameworthy act. Alicke’s (2008) description of blameworthiness is representative: “A blameworthy act occurs when an actor intentionally, negligently, or recklessly causes foreseen, or foreseeable, harmful consequences without any compelling mitigating or extenuating circumstances” (p. 179). Although these experts agree about the kinds of acts that elicit blame judgments, arguably the most recurrent disagreement is the extent to which blame judgments are shaped by epistemic forces (i.e., individuals’ knowledge and reasoning) or motivational forces (i.e., individuals’ desires and goals). Some researchers accord inferential reasoning a central role in shaping moral judgments, whereas other researchers assign more weight to motivated cognition in shaping such judgments.

For example, the path model of blame (Guglielmo et al., 2009; Malle et al., 2014) accords less importance to motivational factors and, instead, delineates a specific cognitive structure (or “path”) that individuals follow to produce blame judgments. What makes judgments moral is that they are directed at agents who are presumed, accused, or shown to have performed behaviors that caused or permitted harm to occur. Within this structure, blame judgments involve integrating information relevant to certain critical concepts and “testing” whether the criteria are met in a graded and systematic manner. These criteria include *intentionality* (whether the agent brought about the event intentionally), *obligation* (whether the agent should have prevented the norm-violating event), and *capacity* (whether the agent could have prevented the event). If the agent is judged to have acted intentionally, then the perceiver considers the agent’s reasons for acting. Depending on the justification that these reasons provide, these inputs give rise to a dichotomous ‘yes-or-no’ and a scalar ‘more-or-less’ judgment of blame—minimal blame if the agent was justified in acting this way or maximal blame if the agent was not justified in acting this way.

In contrast, motivated blame models suggest that the consideration of causal and mental information is secondary to and biased by early-emerging moral judgments and by a general desire to blame. The various models share the view that moral judgments quickly emerge in response to a norm violation, and perceivers consider the details of the event (e.g., intentionality) later, often as a post-hoc rationalization of their judgements (Alicke, 2000; Pettit & Knobe, 2009). From an attributionist perspective, the general desire to assign blame stems from a motivation to maintain a sense of predictability and control. That is, attribution processes are not only a means of providing an individual with a veridical view of his/her world, but a means of forming and maintaining one’s effective exercise of control in the world (e.g., Kelley, 1971). This idea is related to the view that moral judgment is largely motivated by the need to evaluate other individuals, which allows people to navigate their social

environment in more adaptive ways (Pizarro & Tannenbaum, 2012; Tannenbaum et al., 2011).

Indeed, blaming or other expressions of outrage may allow individuals to navigate their social environment in adaptive ways if such expressions also signal positive information about oneself. For example, Jordan et al. (2016b) tested the notion that punishing a wrongdoer may signal one's virtue to observers. In this work, one set of participants, the Signalers, were willing to sacrifice their own money and pay to punish a third party for selfish behavior—even though the Signalers themselves had not been personally mistreated. A second group of participants was more likely to trust the Signalers who punished selfishness and they were right to be trusting, as Signalers actually behaved in a more trustworthy manner. More recent studies by Jordan and Kteily (2020) found that participants expected to be evaluated more positively (by Deciders/perceivers) regardless of if they punished an unambiguous wrongdoing (very severe and likely to be true) or an ambiguous wrongdoing (less severe and less likely to be true), and that participants punished at higher rates when their behavior was public. Together, these experiments underscore a “virtue-signaling” account of expressing moral outrage, such that the costs of expressing outrage may be redeemed by the social benefits of signaling one's trustworthiness and bolstering one's reputation. These studies suggest that expressing outrage is motivated by the increased ability to navigate one's social environments. This work also implies that there is a social cost to *not* displaying outrage, such that individuals are not viewed by others as trustworthy. Indeed, recent work demonstrates that individuals who decide to forgive wrongdoers, rather than punish them, are perceived as blameworthy and as having bad moral character (Gardner & Monroe, 2018).

However, the act of blaming also carries substantial social costs for blamers, such as retaliation towards the blamer by the norm-violator or reputational damage for the blamer when an accusation is unsupported or unwarranted (e.g., MacCoun, 2005; Mikula et al., 1998). As such, the socially-regulated blame perspective (Monroe & Malle, 2019) posits that the norms of moral criticism demand that individuals have warrant when they blame others. Together, the requirement for warrant and the potential social costs of blaming (for blamers) motivate individuals to carefully process available blame-relevant information, such as an agent's causality and other mental states (noted above). A series of studies by Monroe and Malle (2019) found that moral perceivers do systematically grade and update blame judgements in response to the strength of new causal and mental information about an agent's intentionality, reasons, and preventability. For example, relative to initial blame for a violation whose intentionality is ambiguous, participants increased blame when they learned that the violation was intentional

and decreased blame when they learned that the violation was unintentional. The authors also found that when social pressure to form fair and justified moral judgments was reduced, participants' blame judgments were no longer systematic and, instead, showed a bias towards over-blaming. Overall, the social costs of blaming motivate individuals to calibrate their judgements to the causal and mental information surrounding a moral violation.

### The link between blaming and cognitive costs

The above-reviewed literature underscores the role of social costs in motivating outrage and blame, and it suggests that people are less motivated to systematically update moral judgements of blame when such costs are absent. Most notably, Monroe and Malle (2019) found that systematic and graded blame updating occurred even when participants' cognitive resources were successfully reduced. While this finding suggests that cognitive distractions or a lack of motivation to engage in effortful cognition do not deter individuals from processing incoming morally-relevant information and altering blame decisions in a calibrated manner, what is less understood is the degree to which individuals' choices to express outrage and blame is linked to the inherent cognitive costs of doing so. This question is relevant in modern digital contexts, where respondents are more frequently exposed to outrage-inducing stimuli, which may be fatiguing, while at the same time they are more readily able to express outrage that can potentially reach millions (Brady et al., 2017; Crockett, 2017). Deeper tendencies for motivated blame may be related to how cognitively taxing the act of blaming feels.

On average, individuals prefer to avoid effort (Hull, 1943; Kool & Botvinick, 2018). There is an intrinsic and subjective cost associated with effort, such that individuals will avoid or disengage from activities that are cognitively demanding in order to avoid or reduce the inherent price of effort exertion (Kool & Botvinick, 2013, Kool & Botvinick, 2014; but see Inzlicht et al., 2018). Kool et al. (2010) have provided direct empirical support for this idea using a behavioral paradigm, called the Demand Selection Task (DST). In this task, participants face a recurring choice between two alternative lines of action, each associated with different levels of cognitive demand, and they consistently show a preference for the less effortful course of action (Kool et al., 2010). Previous accounts of mental effort have linked effort with emotional aversion (e.g., Inzlicht et al., 2015; Kurzban, 2016). In addition to effort and aversion, cognitive costs can be conceptualized as how much work is extracted from a person for a given level of achievement (Hsu et al., 2017), such that a person feels less efficacious when a task requires a lot from them in order to attain a certain goal or achievement level. Thus, in the current work, we examined felt effort, aversion, and inefficacy as aspects of cognitive costs.

## Current studies: the blame selection task

Recent work suggests that individuals do in fact avoid cognitive work in moral and social contexts: People choose to avoid empathy for strangers, a preference that is linked to felt cognitive costs (Cameron et al., 2019). As its own moral emotion, the choice to feel and express outrage via blaming may be sensitive to whether it is more cognitively costly than, for example, engaging in empathy for wrongdoers or morally disengaging (being objective). In particular, the choice to blame may be sensitive to one's desire to make the most accurate or efficacious moral decision, stemming from social demand (e.g., rewards, norms, etc.; Jordan et al., 2016b; Kundu & Cummins, 2013) or from the fact that some moral decisions are not intuitive and, instead, require more deliberation from the perceiver (e.g., moral dilemmas; Kahane et al., 2012).

These points have the potential to be quite important for the field of moral psychology, particularly in the face of the socially-regulated blame perspective, which posits that the repeated requirement for a blame warrant makes the processing of morally-pertinent information fast and effortless. Within moral psychology, empirical arguments for and against motivated accounts of blaming depend largely on how researchers interpret participants' judgments of agent blameworthiness, causation, intent, desire, foresight, and other factors that legal models identify as related to blame (see Ames & Fiske, 2015). Further, the typical approach is to present participants with moral cases and to subsequently assess their ascriptions of blame via self-reports, without providing participants with the actual *choice* of morally engaging or disengaging in real time. This methodological innovation is important because revealed choices towards generating outrage and blaming (or not doing so) could be more consequential—and thus, potentially more costly—than self-reports of blame intentions or attitudes. To our knowledge, only one other study has directly examined choices to engage in blame (Ames & Fiske, 2015): Participants were more likely to select a blaming task from among five options after reading about a single intentional (versus unintentional) harm. The current set of studies expand beyond this work by examining choices to engage in blaming over a series of instances, and critically, examining the associations between blaming and felt cognitive costs.

To accomplish these objectives, we developed the *Blame Selection Task* (BST), where respondents make a series of choices between engaging in blame and an alternative action. Adapted from the logic of the DST (Kool et al., 2010) and Empathy Selection Task (Cameron et al., 2019), this approach allows us to explore if respondents' baseline motivation is to feel outrage and blame individuals when presented with moral transgressions. Individuals are well familiar with managing moral outrage; in fact, there

has been much colloquial discussion of how individuals selectively choose to expose themselves to news media in order to amplify or diminish feelings of outrage (e.g., Ley, 2017; Mills, 2017). With the BST, we examine motivations towards outrage and blame through spontaneous choices, and with measures of cognitive costs, we examine whether such moralizing feels effortless or effortful.

Across studies, there were a number of procedural variations in order to test the generalizability of our results (Yarkoni, 2019). First, in all studies, participants chose between two card decks, one of which was always a "BLAME" deck, over repeated trials. Each deck entailed different instructions on how to respond to the moral transgressions that are presented. The nature of the contrast deck differed across studies. In some studies, blaming the transgressor was contrasted against morally disengaging by objectively describing the transgressor's actions (Studies 1, 2, 4a, 5a, 5b, 6, and 7); and in other studies, blaming the transgressor was contrasted against morally engaging by empathizing with the transgressor's inner experiences (Studies 3 and 4b). We changed the contrast deck in order to examine whether blame selection would be comparable even as competing courses of action changed (i.e., objective detachment or empathy for the transgressor). Second, we examined blame selection and cognitive costs in response to realistic moral transgressions that are depicted in more than one format (e.g., vignettes versus images versus images with captions). Whereas verbal descriptions of transgressions are often quite clear, images of transgressions may be more equivocal and difficult to interpret, thus leading to changes in blame choice behavior. Alternatively, images may be more effective than words in motivating onlookers' behaviors (e.g., Winkielman & Gogolushko, 2018), thus leading to increased blame selection. Third, we examined whether blame selection shifted depending on whether respondents viewed moral transgressions before making their choices (Studies 1, 2, 3, 5a, 6, and 7) or after making their choices (Studies 4a, 4b, and 5b), which allowed us to differentiate context-bound motivations to blame (in response to particular moral acts) from context-general motivations to blame. These variations are detailed both below and in the Supplementary Materials.

Our dependent variables were the choice to blame (i.e., the proportion of trials where outrage and blame are selected, compared against chance—50%) and the perceived cognitive costs (i.e., effort, aversion, and inefficacy) of each deck (the BLAME deck and the alternative deck). Across studies, we explored whether the proportion of trials that the BLAME deck was chosen significantly varied from chance, specifically whether the patterns of blame choice differed depending on the contrast deck used (objectivity or empathy), the type of moral stimuli presented (vignettes or images), and the position of the decks (deck choices before or after the

moral stimuli). In keeping with the above-reviewed literature on the avoidance of cognitive work (e.g., Kool & Botvinick, 2013, 2014, 2018; Kool et al., 2010), even within in socio-moral contexts (e.g., Cameron et al., 2019), we hypothesized that respondents' choice to generate outrage and blame others would negatively relate to the degree that blaming (relative to the alternative deck) felt cognitively costly (assessed via ratings of effort, aversion, and inefficacy).

## General methods, measures, and procedures

### Participants

All studies were approved by the Research Ethics Board at the University of Toronto. Participants were voluntarily recruited for 30 min of participation on TurkPrime (Litman et al., 2016), an online data collection platform that recruits diverse research samples. All responses provided by participants were strictly anonymous. Those who failed attention checks concealed within the measures of individual differences or who failed to follow the writing prompts (e.g., entering incoherent responses or entering responses that did not match the instructions for the deck) were excluded from analyses, as they were not engaged with the task. According to a power analysis using the average effect in social psychology ( $d=0.40$ ; Richard et al., 2003), with a one-sample  $t$ -test design, we could achieve 80% power in a two-tailed test with 52 participants and 90% power in a two-tailed test with 68 participants. Study 1 included 102 participants; Study 2 included 110 participants; Study 3 included 201 participants; Study 4a included 205 participants; Study 4b included 203 participants; Study 5a included 219 participants; Study 5b included 205 participants; Study 6 included 210 participants; and Study 7 included 217 participants. The sample sizes in Studies 3 to 7 increased from Studies 1 and 2 in order to examine individual difference correlations; power analyses suggested a minimum sample size of 193 to detect a modest correlation ( $r=0.20$ ) with 80% power in a two-tailed test. Table S1 in the Supplementary Materials presents the demographic details for each sample, including the number of respondents excluded in each study.

### The Blame Selection Task

The BST was programmed on Qualtrics survey software (Qualtrics, Provo, UT). In all studies, participants chose between two card decks, one of which was always a "BLAME" deck, over repeated trials. Deck positioning was counterbalanced in Studies 1 to 6; two separate Qualtrics surveys were run sequentially, such that roughly half of the study sample completed the BST with the BLAME deck on the left or right side of the task, and the rest of the study

sample subsequently completed the BST with the BLAME deck on the opposite side (see Table S2 in the Supplementary Materials for exact  $n$ 's in each study). Given the differences in moral stimuli, deck choices, and deck position between studies, the pre-task and trial-level instructions of the BST varied slightly across studies and are described both below and in the Supplementary Materials. All studies first included a sample trial after the pre-task instructions to demonstrate to participants the type of moral stimuli and the decks that they would see, as well as to demonstrate sample responses for the BLAME and alternative decks. All participants then completed two practice trials to familiarize themselves with the sequence of a trial. The moral stimuli (i.e., vignettes or images) were presented in a fixed random order (see the Supplementary Materials for more details as related to each study). While there was no time limit for participants to choose to blame or not, a timer was implemented so that participants could not submit the written response until after a certain amount of time had elapsed (Studies 1 to 4b = 15 s; Studies 5a to 6 = 8 s; Study 7 = 6 s).

### Post-task assessments of cognitive costs and individual differences

#### Cognitive costs

Immediately after the BST, participants across all studies completed items adapted from the NASA Task Load Index (Hart & Staveland, 1988), which we used to assess how cognitively costly they found each deck (i.e., BLAME and either DESCRIBE or FEEL). For each deck, participants answered four questions: "How mentally demanding was this deck?", "How hard did you have to work to accomplish your level of performance with this deck?", "How insecure, discouraged, irritated, stressed, and annoyed were you by this deck?", and "How successful were you in accomplishing what you were asked to do in this deck?" (1 = *Very Low* to 5 = *Very High*). For each deck, the first two items were averaged into an index of perceived effort. The third and fourth items correspond to perceived aversion and perceived efficacy, respectively. All three indices contributed to one's assessment of how cognitively costly each deck felt.

#### Measures of individual differences

Across studies, respondents completed a number of individual difference measures. First, we included three measures of moral personality, as a means to test the convergent validity of the BST. In Studies 1 to 7, the Moral Convictions Scale (Skitka, 2010) assessed the degree to which participants exhibited strong moral convictions about the transgressions they read/saw. Given that individuals with stronger convictions about moral issues prefer greater social and physical

distance from attitudinally divergent others and endorse greater intolerance towards attitudinally divergent others (Skitka et al., 2005), it is expected that participants who more strongly moralize wrongdoings are more motivated to blame transgressors. In Studies 2 to 7, the Moral Identity Scale (Aquino & Reed, 2002) assessed self-importance of moral identity. Given the links between moral identity and increased pro-social behaviors and moral emotions (e.g., Aquino & Reed, 2002; Hardy, 2006; Stets & Carter, 2006), participants who tie morality more closely to their self-concept might choose to blame transgressors more often. In Study 7, the Blame Intensity Inventory (Gill & Cerce, 2021) assessed individual differences in the propensity to experience intense blame reactions; it is expected that participants with a greater tendency to have hostile affective reactions towards offenders would choose to blame transgressors more frequently. Outside of moral personality, we assessed *descriptive* and *injunctive social norms*, as well as *personal norms*, about blaming in Studies 2 to 7 (drawing on five items from Cameron et al., 2019). Recent work (Son et al., 2019) finds that victims and jurors will reference group preferences to inform their own punitive decisions, becoming increasingly punitive as groups express a desire to punish. Thus, to the degree that participants believe that others choose blame, that others value blame, and to the degree that participants themselves view blaming as desirable, they might select into blame-eliciting contexts. Finally, the Empathy Index (Jordan et al., 2016a) assessed trait empathy in Studies 1 to 7, as it might be expected that highly empathetic participants would have a higher blame motivation if they are empathizing with the victim(s) of the transgression (e.g., see Leliveld et al., 2012 on the relationship between observers' empathy and punishment of offenders). Details about each measure, including how items were rated, sample items, subscales, as well as the Cronbach's  $\alpha$  values (Table S3) and mean scores (Table S4) across all studies, are found in the Supplementary Materials. Participants in Studies 2 to 6 also completed the Ten-Item Personality Inventory (Gosling et al., 2003), but results pertaining to this measure were not of central interest in the current paper and are presented in Table S5 in the Supplementary Materials.

## Data analyses

Across all studies, some of the blame choice variables were non-normally distributed (i.e., possessing a slight positive or negative skew). As such, we conducted both non-parametric and parametric tests to examine the degree to which participants chose to blame compared to chance (one-sample tests), as well as the blame choice patterns based on the ambiguity of the moral transgressions and on the type of

moral stimuli presented (paired sample tests; described further below). It is noted that the pattern of results remained the same in all studies. We present results of non-parametric Wilcoxon tests within the main text and parametric *t*-test statistics in the Supplementary Materials. All other *t*-test assumptions, regarding the scale of measurement, random sampling, adequacy of sample size, and equality of variance in standard deviation, were met.

Across studies that used moral vignettes (Studies 1 to 4b) and moral images (Studies 5a to 6), non-parametric independent samples tests found that blame choice (i.e., the proportion of trials where participants chose the BLAME deck) of the participants who saw the BLAME deck on the left side of the BST did not significantly differ from the blame choice of the participants who saw the BLAME deck on the right side, which confirms that there was no left- or right-side bias for selecting the BLAME deck. See Table S2 in the Supplementary Materials for test statistics and effect sizes across all studies (vignettes and images).

## Studies 1 to 4b: blame selection task with moral vignettes

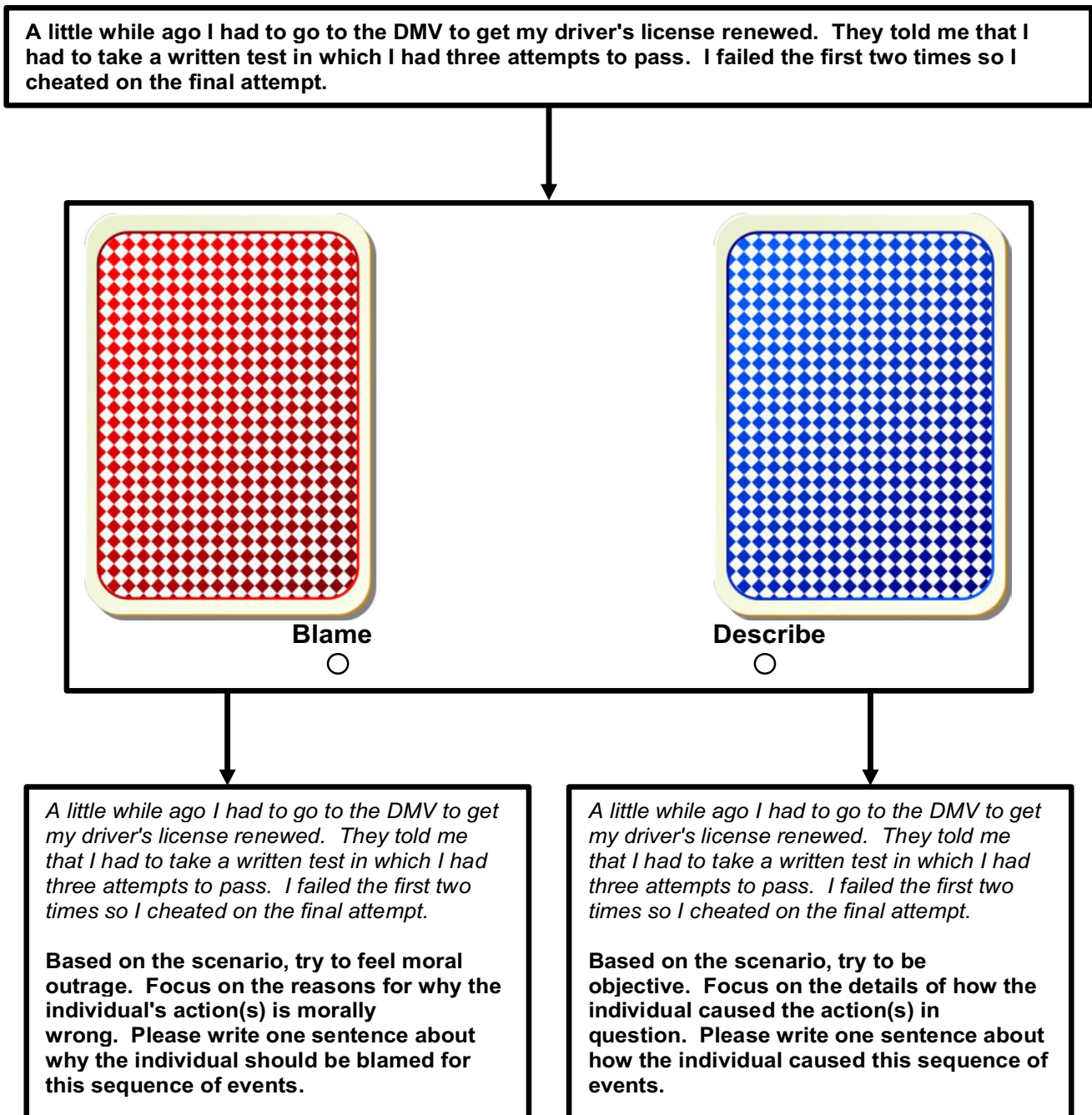
Thirty-four moral vignettes (brief written descriptions of a target individual engaging in a real-life action(s) that is morally wrong e.g., assault, robbery, infidelity, etc.) were obtained from a moral vignettes database by Knutson et al. (2010). Six vignettes were obtained from the supplementary materials published by Gill and Getty (2016) and slightly edited. We then piloted these 40 vignettes, which were written in first-person, with 109 participants, who rated each vignette on indices of *Harm* (low Care) and *Moral Appropriateness*.

Of the 40 vignettes, 26 were included in Study 1. The full procedure for selecting the 26 moral vignettes in Study 1 and the average Care and Moral Appropriateness rating for each vignette (Table S6) are in the Supplementary Materials. In Study 2 and subsequent studies that used vignettes, we explored whether respondents were more likely to blame in response to certain types of moral transgressions than others. Whereas Jordan and Kteily (2020) illustrate the role of perceived reputational incentives on participants' punishment of unambiguous and ambiguous acts, we sought to explore respondents' first-person management of their choice to blame immoral acts of varying levels of severity. Based on a composite (average score) of the Care and Moral Appropriateness ratings of the 40 moral vignettes that were piloted prior to Study 1, 16 vignettes were grouped and labelled as *clear moral transgressions*, where the action(s) was conspicuously harmful and morally inappropriate, and 16 vignettes were grouped and labelled as *ambiguous moral transgressions*,

where the action(s) was less conspicuously harmful and morally inappropriate. See the Supplementary Materials for further details and statistical justification for this grouping procedure. Mean Care and Moral Appropriateness ratings and the composite Care-Moral Appropriateness scores of the 32 vignettes used in Studies 2, 3, 4a, 4b, and 7 are in Table S7.

### Study 1 methods

In Study 1, participants ( $N=102$ ) were instructed that they would complete a series of trials in which they would read a first-person vignette of an individual committing an action and then see two decks of cards to choose between freely.



**Fig. 1** Schematic example of blame selection task trial (Studies 1 to 3). *Note.* Over repeated trials, respondents choose a deck after reading a first-person vignette of an individual committing an action. Based upon choice, respondents are instructed to feel moral outrage or to be objective and make a written response. The schematic includes a

target stimulus used in Studies 1, 2, 3, 4a, and 4b (from the moral vignettes database; Knutson et al., 2010). See main text for modified variants in Studies 2, 3, 4a, 4b, and 7, which also used moral vignettes

Participants were told that if they chose the BLAME deck, they would be instructed to feel moral outrage, focus on the moral details of the individual and the action(s) committed, and write one sentence about why the individual should be blamed for the sequence of events. If participants chose the objective (“DESCRIBE”) deck, they would be instructed to be objectively detached, focus on the features of the individual and the action(s) committed, and write one sentence about how the individual caused the sequence of events (see the Supplementary Materials for the pre-task instructions of all studies). Participants then completed 26

trials of the BST. If they chose the BLAME deck, they saw the following prompt: “Based on the scenario, try to feel moral outrage. Focus on the reasons for why the individual’s action(s) is morally wrong. Please write one sentence about why the individual should be blamed for this sequence of events”. If participants chose the DESCRIBE deck, they saw the following prompt: “Based on the scenario, try to be objective. Focus on the details of how the individual caused the action(s) in question. Please write one sentence about how the individual caused this sequence of events”. Figure 1 depicts a single trial of the BST in Study 1.

**Table 1** Median Blame Choice and Mean NASA Task Load Index Ratings for BLAME and Alternative decks (Studies 1 to 7)

Median	Study 1	Study 2	Study 3*	Study 4a	Study 4b*	Study 5a	Study 5b	Study 6	Study 7
<i>Blame Choice:</i>	0.58	0.66	0.66	0.63	0.63	0.13	0.10	0.43	0.20
<i>Z</i>	1.55	4.44	8.63	2.45	3.60	−11.88	−10.80	−6.21	−10.28
<i>p</i>	.120	<.001	<.001	.007	<.001	<.001	<.001	<.001	<.001
<i>r<sup>a</sup></i>	0.15	0.42	0.61	0.17	0.25	−0.80	−0.75	−0.43	−0.70
<i>Blame Choice:</i>	–	–	–	–	–	–	–	–	0.53
<i>Z</i>									0.67
<i>p</i>									.505
<i>r</i>									0.05
M (SD)	Study 1	Study 2	Study 3*	Study 4a	Study 4b*	Study 5a	Study 5b	Study 6	Study 7
<i>Perceived Effort:</i>									
BLAME	3.38 (1.18)	3.47 (1.07)	3.30 (1.07)	3.43 (1.14)	3.24 (1.16)	3.80 (1.02)	3.76 (0.99)	3.59 (1.07)	3.54 (1.11)
OTHER deck	3.14 (1.16)	3.46 (1.12)	3.44 (0.97)	3.37 (1.20)	3.73 (0.99)	2.88 (1.24)	3.20 (1.08)	3.05 (1.18)	3.29 (1.07)
<i>t (df)</i>	1.55 (101)	.12 (109)	−2.01 (200)	.60 (204)	−5.46 (202)	9.86 (218)	6.83 (204)	6.69 (209)	3.45 (216)
<i>p</i>	.125	.904	.046	.547	<.001	<.001	<.001	<.001	.001
Cohen’s <i>d<sup>a</sup></i>	0.16	0.01	−0.14	0.04	−0.38	0.67	0.47	0.46	0.24
<i>Perceived Aversion:</i>									
BLAME	2.87 (1.44)	2.85 (1.32)	2.83 (1.25)	2.86 (1.37)	2.90 (1.29)	3.17 (1.36)	3.27 (1.39)	2.94 (1.32)	3.14 (1.26)
OTHER deck	2.66 (1.35)	2.63 (1.36)	2.62 (1.13)	2.80 (1.39)	3.17 (1.27)	2.27 (1.25)	2.56 (1.34)	2.35 (1.31)	2.60 (1.15)
<i>t (df)</i>	1.38 (101)	1.68 (109)	2.71 (200)	.62 (204)	−2.71 (202)	8.79 (218)	6.84 (204)	6.53 (209)	7.48 (216)
<i>p</i>	.172	.048	.007	.535	.007	<.001	<.001	<.001	<.001
Cohen’s <i>d</i>	0.14	0.16	0.19	0.04	−0.19	0.60	0.48	0.45	0.50
<i>Perceived Efficacy:</i>									
BLAME	3.76 (1.08)	3.86 (1.04)	3.96 (0.87)	3.86 (1.01)	3.97 (0.84)	3.40 (1.25)	3.45 (1.27)	3.79 (1.09)	3.61 (1.04)
OTHER deck	3.70 (1.09)	3.83 (1.00)	3.84 (0.88)	3.54 (1.23)	3.46 (1.08)	4.22 (0.92)	4.15 (0.89)	4.06 (0.91)	4.18 (0.69)
<i>t (df)</i>	.48 (101)	.37 (109)	2.02 (200)	3.26 (204)	5.83 (202)	−8.88 (218)	−7.30 (204)	−3.23 (209)	−8.27 (216)
<i>p</i>	.634	.712	.045	.001	<.001	<.001	<.001	.001	<.001
Cohen’s <i>d</i>	0.05	0.04	0.14	0.23	0.41	−0.60	−0.51	−0.23	−0.56

Study 7 presents blame choice scores for both image-based (top) and vignette-based (bottom) transgressions. Studies 3 and 4b used the FEEL deck as the alternative choice to the BLAME deck. Cognitive cost differences for Studies 3 and 7 were based on the average perceived effort, perceived aversion, and perceived efficacy scores of the BLAME and alternative (FEEL or DESCRIBE) decks across both administrations of the NASA Task Load Index

<sup>a</sup>The effect size *r* and Cohen’s *d* values are specific to within-subjects effects



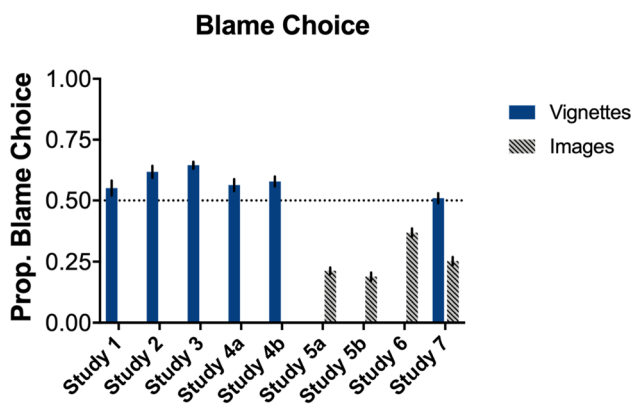


Fig. 2 Blame Choice in Studies 1 to 7 (Vignette-Based and Image-Based Studies). Note. Error bars reflect standard error of the mean

### Study 1 results

#### Blame choice

Table 1 presents the descriptive and inferential statistics (medians, one-sample Wilcoxon signed-rank tests, and effect size  $r^1$ ) for blame choice across all studies. Figure 2 depicts blame choice across all studies. In Study 1, although respondents chose to blame more than describe, a one-sample Wilcoxon signed-rank test found that blame choice did not significantly differ from chance (50%),  $Z = 1.55, p = .120$ , with a small blame choice effect observed ( $r = .15$ ).

#### Cognitive costs

Table 1 also presents average ratings of perceived effort, perceived aversion, and perceived efficacy for the BLAME and alternative decks (i.e., DESCRIBE or FEEL) across all studies.<sup>2</sup> We report differences in these ratings between Blame and Objectivity or Empathy in each study. Paired sample  $t$ -tests revealed that differences between the BLAME and DESCRIBE decks in felt effort,  $t(101) = 1.55, p = .125$ , 95% CI of  $M_{\text{difference}} [-.07, .54], d = 0.16$ , aversion,  $t(101) = 1.38, p = .172$ , 95% CI of  $M_{\text{difference}} [-.10, .53], d = 0.14$ , and efficacy,  $t(101) = .48, p = .634$ , 95% CI of  $M_{\text{difference}} [-.22, .35], d = 0.05$ , were small and not significant in Study 1. Together, these non-significant differences in cognitive costs could account for why there was no overall motivation among respondents in this study to choose to blame wrongdoers.

<sup>1</sup> The interpretation values for  $r$  commonly reported in published literature are: 0.10–<0.3 (small effect), 0.30–<0.5 (moderate effect) and  $\geq 0.5$  (large effect).

<sup>2</sup> Because the NASA Task Load index was administered twice in Studies 3 and 7 (i.e., halfway through the BST and at end of the BST), we computed the average effort, aversion, and efficacy scores between both time-points for the BLAME and FEEL/DESCRIBE decks.

### Study 2 methods: refinement of Study 1 BST

In Study 2, participants ( $N = 110$ ) completed a BST equivalent to what is described in Study 1. However, trial-level instructions of the BLAME deck were slightly revised to ensure that the intensity associated with the term “outrage” was not inadvertently biasing respondents’ deck choices: “Based on the scenario, try to feel moral *disapproval*. Focus on the reasons for why the individual’s action(s) is morally wrong. Please write one sentence about why the individual should be blamed for this sequence of events” (italics added to indicate revision).

### Study 2 results

#### Blame choice

In contrast to Study 1, blame choice in Study 2 was significantly greater than chance,  $Z = 4.44, p < .001$ , with a moderate blame choice effect observed among participants ( $r = .42$ ).

#### Cognitive costs

With respect to cognitive costs differences in Study 2, the BLAME deck felt significantly more aversive than the DESCRIBE deck,  $t(109) = 1.68, p = .048$ , 95% CI of  $M_{\text{difference}} [.04, .48], d = 0.16$ , but felt no different in effort,  $t(109) = .12, p = .904$ , 95% CI of  $M_{\text{difference}} [-.21, .24], d = 0.01$ , and efficacy,  $t(109) = .37, p = .712$ , 95% CI of  $M_{\text{difference}} [-.16, .23], d = 0.04$ .

### Study 3 methods: blocking vignettes by moral ambiguity with empathy contrast

In Study 3, the 32 vignettes were blocked based on moral ambiguity of the transgression, which allowed us to examine participants’ management of outrage and blame in response moral acts of differing levels of severity. Some respondents ( $n = 100$ ) read all 16 clear moral transgressions first and other respondents ( $n = 101$ ) read all 16 ambiguous moral transgressions first. Unlike in Study 2, the contrast deck to blame was a “FEEL” deck (similar to the empathy deck used in Cameron et al., 2019), which allowed us to examine whether blame choice would change as the alternative action changed (i.e., from objective detachment to empathy). In the pre-task instructions, respondents were told that if they chose the FEEL deck, they would be instructed to have empathy and write one sentence about the feelings and experiences of the individual who committed the action(s). Trial-level instructions for the BLAME deck were revised in this study to ensure congruency with the FEEL deck in the word-count and

syntax (see the Supplementary Materials). If participants chose the BLAME deck, they saw the following prompt: “Based on the scenario, try to feel moral outrage toward the individual who committed the action(s). Morally focus on the blameworthiness of this individual’s actions(s). Please write one sentence about why the individual, who committed the action(s), should be blamed”. If participants chose the FEEL deck, they saw the following prompt: “Based on the scenario, try to feel what the individual who committed the action(s) feels. Empathically focus on this individual’s internal emotional experiences. Please write one sentence about what the individual, who committed the action(s), feels and experiences”. Finally, respondents in Study 3 completed to the NASA Task Load Index twice, once after the first half of trials (“Please answer the following questions about the [BLAME or FEEL] deck, thus far”) and once after the second half of trials (“Please answer the following questions about the [BLAME or FEEL] deck, based on your experience during the second-half of the task”), in order to collect cognitive costs ratings separately for clear and ambiguous moral transgressions.

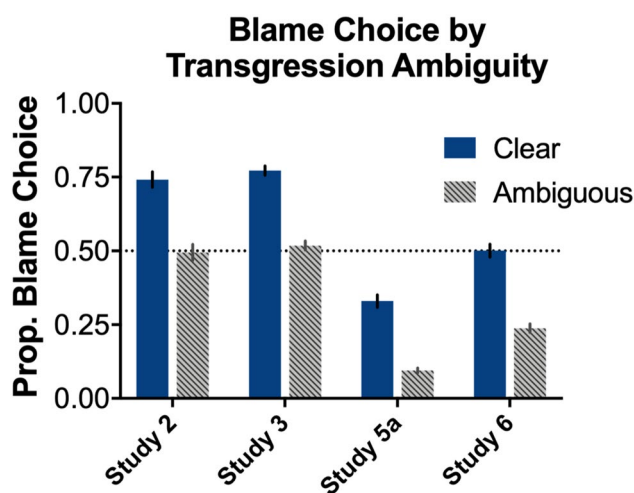
## Study 3 results

### Blame choice

Similar to Study 2, blame choice was significantly greater than chance in Study 3,  $Z=8.63$ ,  $p<.001$ ,  $r=0.61$ . Between Studies 2 and 3, there was no significant difference in blame choice, as evidenced by a Wilcoxon rank-sum test,  $Z=.62$ ,  $p=.537$ ,  $r=.04$ . Together, these results suggest that, irrespective of the alternative deck choice (either objectivity or empathy towards transgressors), respondents displayed a moderate-to-strong motivation to feel outrage and blame individuals depicted in the moral vignettes after reading the vignettes.

### Blame choice by moral ambiguity of vignettes

We computed a blame choice score (i.e., the proportion of trials that the respondent chose the BLAME deck) for the groups of clear and ambiguous moral transgressions for each participant in Studies 2 and 3, where vignettes—labelled as either clear or ambiguous—preceded deck choice. A Wilcoxon signed-rank test was used to compare blame choice between the clear and ambiguous moral transgressions in all relevant studies. The standardized statistic for this test is based on differences between blame choice for the two levels of moral ambiguity, specifically when blame choice for ambiguous moral transgressions is subtracted from blame choice for clear moral transgressions. Figure 3 depicts blame choice by the moral ambiguity of transgressions (i.e., clear versus ambiguous) across all relevant studies.



**Fig. 3** Blame Choice by Moral Ambiguity of Transgressions in Studies 2, 3, 5a, and 6. *Note.* Error bars reflect standard error of the mean

As evidenced by paired samples Wilcoxon signed-rank tests, blame choice was significantly greater for the clear moral transgressions than for the ambiguous moral transgressions within both studies (Study 2:  $Z=8.13$ ,  $p<.001$ ,  $r=.78$ ; Study 3:  $Z=11.66$ ,  $p<.001$ ,  $r=.82$ ), with a very robust effect of transgression ambiguity on participants’ blame choices. One-sample Wilcoxon signed-rank tests found that participants chose the BLAME deck at a rate significantly greater than chance (i.e., greater than 50% of trials) after reading the clear moral transgressions in both studies (Study 2:  $Median_{choice}=0.88$ ,  $Z=6.60$ ,  $p<.001$ ,  $r=.63$ ; Study 3:  $Median_{choice}=0.81$ ,  $Z=10.36$ ,  $p<.001$ ,  $r=.73$ ), but did not show a preference for selecting the BLAME deck after reading the ambiguous moral transgressions (Study 2:  $Median_{choice}=0.44$ ,  $Z=-.32$ ,  $p=.746$ ,  $r=-.03$ ; Study 3:  $Median_{choice}=0.50$ ,  $Z=1.27$ ,  $p=.205$ ,  $r=.09$ ). These results suggest that, as opposed to remaining objective or empathizing, respondents chose to blame individuals who commit clearly harmful and morally inappropriate behaviors, but that this blame preference diminished for more ambiguous behaviors.

### Cognitive costs

Consistent with Study 2, the BLAME deck felt significantly more aversive than the FEEL deck in Study 3,  $t(200)=2.71$ ,  $p=.007$ , 95% CI of  $M_{difference}$  [.06, .36],  $d=0.19$ . In Study 3, the BLAME deck also felt significantly more efficacious than the FEEL deck,  $t(200)=2.02$ ,  $p=.045$ , 95% CI of  $M_{difference}$  [.00, .24],  $d=0.14$ . In this study, the BLAME deck felt significantly less effortful than the FEEL deck,  $t(200)=-2.01$ ,  $p=.046$ , 95% CI of  $M_{difference}$  [- .27, - .00],  $d=-0.14$ , which is consistent with recent work that suggests that empathy is cognitively effortful (Cameron et al., 2019).

Notwithstanding these differences in cognitive costs between blaming and empathy, in collapsing across transgression ambiguity, the above analysis might mask differences in cognitive costs based upon the moral ambiguity of the action. Thus, in Study 3, we examined if differences between the BLAME and FEEL decks on perceived effort, aversion, and efficacy varied by the level of moral ambiguity of the action committed.

Table 2 presents the average ratings of perceived effort, perceived aversion, and perceived efficacy for the BLAME and FEEL deck across clear and ambiguous moral transgressions in Study 3. We conducted three 2×2 repeated measures ANOVAs to examine whether differences between BLAME and FEEL decks in effort, aversion, and efficacy were moderated by the moral ambiguity of the transgressions. There was a significant interaction between deck (BLAME, FEEL) and moral ambiguity (clear, ambiguous) on ratings of efficacy,  $F(1, 200) = 9.01, p = .003, \eta_p^2 = .04$ . Pairwise comparisons revealed that that BLAME deck felt significantly more efficacious than the FEEL deck in the block of clear moral transgressions ( $M_{\text{difference}} = .24, SE_{\text{difference}} = .08, p = .003$ ), but not in the block of ambiguous moral transgressions ( $M_{\text{difference}} = .01, SE_{\text{difference}} = .06, p = .938$ ). The interactions between deck and moral ambiguity on ratings of effort,  $F(1, 200) = .91, p = .342, \eta_p^2 = .01$ , and aversion,  $F(1, 200) = 2.14, p = .145, \eta_p^2 = .01$ , were not statistically significant. Although these interactions were non-significant, we examined deck differences in effort and aversion separately by ambiguity level for descriptive completeness. Paired sample *t*-tests revealed that for the clear moral transgressions, the BLAME deck felt significantly less effortful,  $t(200) = -2.08, p = .039, 95\% \text{ CI of } M_{\text{difference}} [-.71, -.02], d = -0.15$ , and more aversive,  $t(200) = 2.90, p = .004, 95\% \text{ CI of } M_{\text{difference}} [.09, .47], d = 0.20$ , than the FEEL deck. However, for the ambiguous moral transgressions, the BLAME and FEEL decks felt no different in effort,  $t(200) = -1.22, p = .222$ ,

95% CI of  $M_{\text{difference}} [-.49, .12], d = -0.09$ , and aversiveness,  $t(200) = 1.58, p = .115, 95\% \text{ CI of } M_{\text{difference}} [-.03, .30], d = 0.11$ . Overall, these results may explain the variable differences in cognitive costs between the BLAME deck and alternative deck in the preceding vignette studies, in which the NASA Task Load Index was administered once and did not differentiate between clear and ambiguous transgression trials.

**Studies 4a and 4b methods: moral vignettes after deck choices**

In Studies 4a ( $N = 205$ ) and 4b ( $N = 203$ ), the card decks preceded the moral vignettes in each trial. Respondents were instructed before the task that they would complete a series of trials where they would first see and freely choose between two decks of cards, after which they would read a first-person vignette of an individual committing an action. This design allowed us to examine context-independent motivations to blame more generally, that is, even without knowing a specific action to blame. The difference between Studies 4a and 4b concerned the contrast deck. In Study 4a, the alternative deck was DESCRIBE (as in Study 2), and in Study 4b, the alternative deck was FEEL (as in Study 3). Study 4a used the same trial-level instructions from Study 2, but with minor changes. If participants chose the BLAME deck, they saw the following prompt: “Based on the scenario, try to feel moral *outrage*. *Focus on the moral details of the action(s) in question*. Please write one sentence to describe why the individual should be blamed”. If respondents chose the DESCRIBE deck, they saw the following prompt: “Based on the scenario, try to be objective. *Focus on the objective details of the action(s) in question*. Please write one sentence to describe how the individual caused the sequence of events” (italics added to indicate revisions). Study 4b used the same trial-level prompts as Study 3. See the Supplementary Material for details.

**Studies 4a and 4b results**

**Blame choice**

Blame choice was significantly greater than chance in Study 4a,  $Z = 2.45, p = .007, r = .17$ , and in Study 4b,  $Z = 3.60, p < .001, r = .25$  (see Table 1; Fig. 2 for more details). Between participants in Studies 4a and 4b, there was no significant difference in blame choice, as evidenced by a Wilcoxon rank-sum test,  $Z = -.15, p = .883, r = -.01$ . Together, these results suggest that, irrespective of the alternative deck choice (either objectivity or empathy towards transgressors), respondents displayed a comparably modest motivation to feel outrage and blame individuals depicted in the moral vignettes before reading the vignettes. That is, respondents

**Table 2** Mean NASA Task Load Index Ratings for BLAME and Alternative decks: Split by Moral Ambiguity of Transgressions in Study 3

<i>M (SD): Study 3</i>	Clear Moral Transgression	Ambiguous Moral Transgression
<i>Perceived Effort:</i>		
BLAME deck	3.31 (1.17)	3.29 (1.13)
FEEL deck	3.49 (1.14)	3.39 (1.05)
<i>Perceived Aversion:</i>		
BLAME deck	2.97 (1.39)	2.69 (1.34)
FEEL deck	2.69 (1.28)	2.55 (1.28)
<i>Perceived Efficacy:</i>		
BLAME deck	4.01 (0.94)	3.91 (0.97)
FEEL deck	3.78 (1.04)	3.91 (0.88)

displayed a general motivation to blame even without knowing the particular context of judgment.

### Cognitive costs

Cognitive cost differences between the decks in Studies 4a and 4b were slightly consistent across the rest of the moral vignette studies (as evidenced by paired sample *t*-tests). In both Studies 4a and 4b, the BLAME deck felt significantly more efficacious than the alternative deck, which is consistent with the patterns found in Study 3; Study 4a:  $t(204) = 3.26, p = .001$ , 95% CI of  $M_{\text{difference}}$  [.13, .52],  $d = 0.23$ ; Study 4b:  $t(202) = 5.83, p < .001$ , 95% CI of  $M_{\text{difference}}$  [.34, .69],  $d = 0.41$ . As with Study 3, the BLAME deck felt significantly less effortful than the FEEL deck in Study 4b,  $t(202) = -5.46, p < .001$ , 95% CI of  $M_{\text{difference}}$  [- .67, - .31],  $d = -0.38$ ; and there were no significant differences in perceived effort between the BLAME and DESCRIBE deck in Study 4a,  $t(204) = .60, p = .547$ , 95% CI of  $M_{\text{difference}}$  [- .14, .26],  $d = 0.04$ . Unlike Study 3, the BLAME deck felt significantly less aversive than the FEEL deck in Study 4b,  $t(202) = -2.71, p = .007$ , 95% CI of  $M_{\text{difference}}$  [- .47, - .07],  $d = -0.19$ ; and there were no significant differences in perceived aversion between the BLAME and DESCRIBE deck in Study 4a,  $t(204) = .62, p = .535$ , 95% CI of  $M_{\text{difference}}$  [- .14, .27],  $d = 0.04$ .

### Studies 5a to 6: Blame Selection Task with moral images

In the next set of studies (Studies 5a to 6), we examined whether blame motivations would differ when moving to a potentially less concrete type of stimulus: moral images. Visual depictions of immoral acts might be more ambiguous and difficult to interpret because there is more information to parse than when presented with a verbal description of such acts. On the other hand, given that bodily responses are more easily engaged by affective information presented with pictures rather than written words (Lang et al., 2005; Larsen et al., 2003; Winkielman & Gogolushko, 2018), respondents may be more compelled to morally engage with visual depictions of moral indiscretions. In Studies 5a and 5b, we examined blame choice in response to images of transgressions, and in Study 6, we added a caption to each image to examine how this might impact blame choice.

A total of 40 moral images were obtained from the Social-Moral Images Database (SMID; Crone et al., 2018) for the BST in Studies 5a, 5b, and 6. Crone et al.'s (2018) normative ratings from 1812 participants were available on eight dimensions, including the level of Care and Moral Appropriateness depicted in the image. Twenty images were grouped and labelled as clear moral transgressions (i.e., the

action in the image was conspicuously harmful and morally wrong), whereas 20 images were grouped and labelled as ambiguous moral transgressions. See the Supplementary Materials for the full procedure and statistical justification for the selection and grouping of the moral images. Average Care and Moral Appropriateness ratings and the composite Care-Moral Appropriateness scores for all 40 moral images are presented in Table S8 (the images are labeled as found in the SMID normative ratings database; Crone et al., 2018).

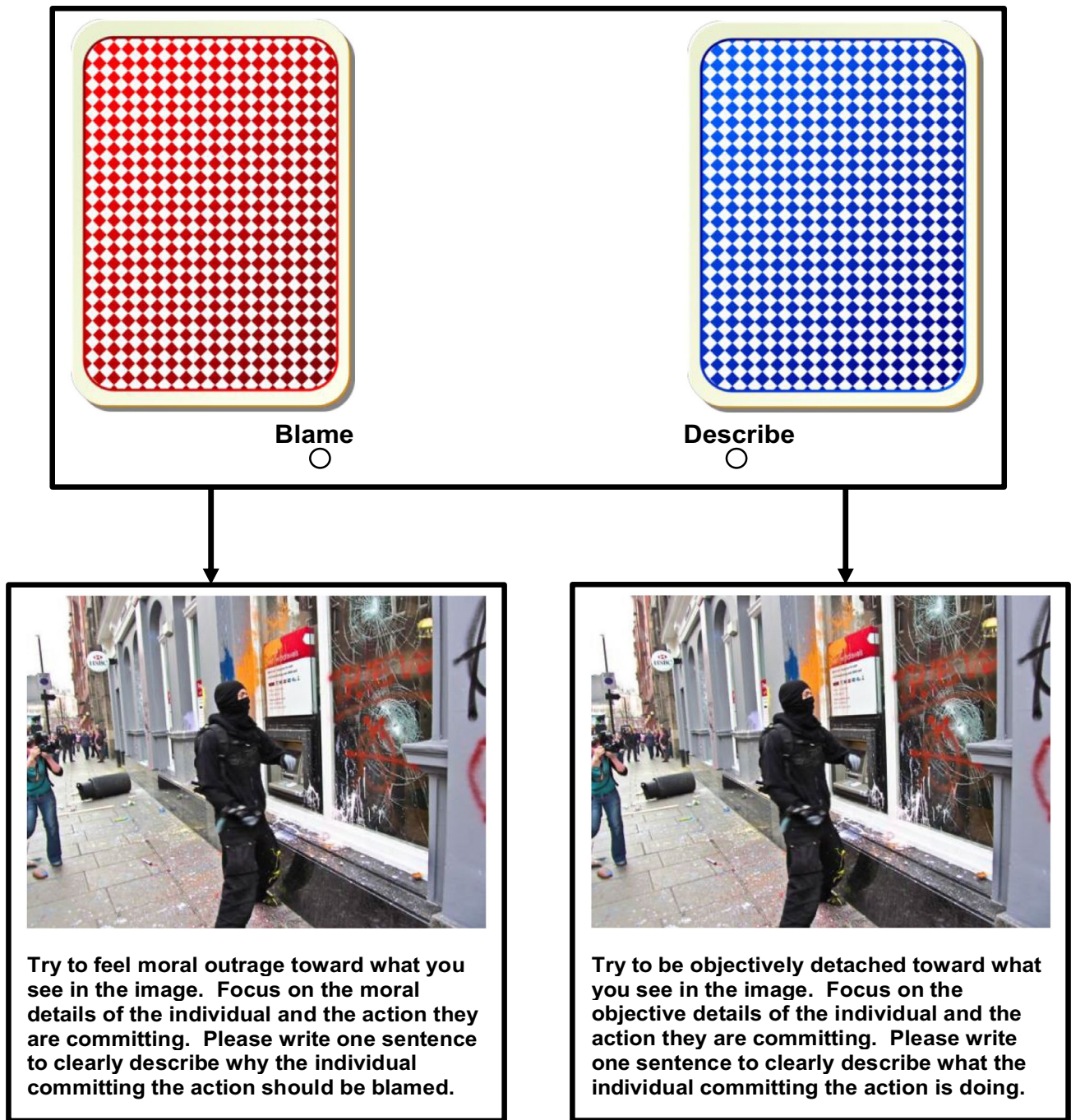
### Studies 5a and 5b methods

The key difference between Studies 5a and 5b is whether the decks (BLAME versus DESCRIBE) were presented after the images (5a) or presented before the images (5b); this was done to parallel the position of moral transgressions in the vignette-based studies. Respondents in Study 5a ( $N = 219$ ) were instructed that they would complete a series of trials where they would first see an image of an individual who is committing an action, after which they would see and freely choose between two decks of cards. Respondents in Study 5b ( $N = 205$ ) were instructed that they would complete a series of trials where they would first see and freely choose between two decks of cards, after which they would see an image of an individual who is committing an action; Figure 4 depicts a single trial of the BST in Study 5b. In both studies, respondents were given the following instructions if they picked the BLAME deck: "Try to feel moral outrage toward what you see in the image. Focus on the moral details of the individual and the action they are committing. Please write one sentence to clearly describe why the individual committing the action should be blamed". If respondents chose the DESCRIBE deck, they were given the following prompt: "Try to be objectively detached toward what you see in the image. Focus on the objective details of the individual and the action they are committing. Please write one sentence to clearly describe what the individual committing the action is doing".

### Studies 5a and 5b results

#### Blame choice

Thus far, the results of the vignette studies suggest that participants show a preference towards outrage and blaming transgressors, rather than descriptively or empathically engaging with them, especially in response to clear moral transgressions. In contrast, one-sample Wilcoxon signed-rank tests revealed that blame choice in response to images (see Table 1 for descriptive statistics) was significantly *less* than chance in Study 5a,  $Z = -11.88, p < .001, r = -.80$ , and in Study 5b,  $Z = -10.80, p < .001, r = -.75$ , which suggests



**Fig. 4** Schematic Example of Blame Selection Task Trial in Study 5b. *Note.* Over repeated trials, respondents choose a deck before seeing an image of an individual committing an action. Based upon choice, respondents are instructed to feel moral outrage or to be objectively detached and make a written response. The schematic includes a target

stimulus used in Studies 5a, 5b, 6, and 7 (from the Social-Moral Image Database; Crone et al., 2018). See main text for modified variants in Studies 5a, 6, and 7, which also used moral images

a strong motivation to *avoid* blaming transgressors when wrongdoings are depicted visually.

**Cognitive costs**

The patterns observed for the cognitive costs of decks in the studies with moral images were different than what was observed in the studies with moral vignettes (see Table 1

for descriptive statistics). The most consistent finding in the studies with moral vignettes was that the BLAME deck felt significantly more efficacious than the alternative deck choice. In contrast, in Studies 5a and 5b, paired sample *t*-tests found that the BLAME deck felt significantly *less* efficacious than the DESCRIBE deck, Study 5a:  $t(218) = -8.88, p < .001, 95\% \text{ CI of } M_{\text{difference}} [-1.00, -.64], d = -0.60$ ; Study 5b:  $t(204) = -7.30, p < .001, 95\% \text{ CI of } M_{\text{difference}} [-.89, -.51], d = -0.51$ . In these studies, the BLAME deck felt significantly *more* effortful than the DESCRIBE deck, Study 5a:  $t(218) = 9.86, p < .001, 95\% \text{ CI of } M_{\text{difference}} [.74, 1.11], d = 0.67$ ; Study 5b:  $t(204) = 6.83, p < .001, 95\% \text{ CI of } M_{\text{difference}} [.40, .72], d = 0.47$ . Finally, the BLAME deck felt significantly *more* aversive than the DESCRIBE deck, Study 5a:  $t(218) = 8.79, p < .001, 95\% \text{ CI of } M_{\text{difference}} [.70, 1.10], d = 0.60$ ; Study 5b:  $t(204) = 6.84, p < .001, 95\% \text{ CI of } M_{\text{difference}} [.51, .92], d = 0.48$ . Moderate-to-large effects of deck choice (BLAME versus DESCRIBE) on felt cognitive costs were observed. These differences in cognitive costs may account for why participants avoided blaming when they were shown images of moral wrongdoings.

## Study 6 methods

In Study 6, respondents ( $N = 210$ ) were instructed that a caption of what the individual was doing would accompany each image before they were to make a choice between the BLAME and DESCRIBE decks. This would allow us to examine if respondents' blame avoidance would diminish if visual moral content was accompanied by verbal information that elucidates the wrongdoing. The captions for the moral images were created by the first author and are listed in the Supplementary Materials. Study 6 used similar trial-level instructions from Studies 5a and 5b, with the addition that respondents were instructed to write the sentence in their own words (i.e., not copy the caption—"Please write one sentence, *in your own words*, to...").

## Study 6 results

### Blame choice

Consistent with Studies 5a and 5b, blame choice in response to images and headlines was significantly less than chance in Study 6,  $Z = -6.21, p < .001, r = -.43$ . Notably, blame choice in response to moral images and captions in Study 6 was higher than blame choice in Studies 5a, where respondents saw only moral images,  $Z = 6.59, p < .001, r = .32$ , which suggests a moderate effect of captions on blame choice differences between study samples. Yet, in tandem with Studies 5a and 5b, the results suggest that people show a motivation to avoid outrage and blame when presented

with visual depictions of transgressions, rather than engage in outrage.

### Blame choice by moral ambiguity of images

As with Studies 2 and 3, blame choice was computed for the groups of clear and ambiguous moral transgressions in Studies 5a and 6, where images preceded deck choices (see Fig. 3). Replicating results from Studies 2 and 3 (which used vignettes), paired samples Wilcoxon signed-rank tests found that blame choice in Study 5a was significantly higher for clear (than ambiguous) moral transgressions,  $Z = 10.89, p < .001, r = .74$ —suggesting that the level of harm and immorality depicted in the pictures largely increases the choice to blame individuals committing such actions.

Similarly, in Study 6, blame choice was significantly greater for clear moral images accompanied by captions than for ambiguous moral images accompanied by captions,  $Z = 11.36, p < .001, r = .78$ . Although this transgression ambiguity difference existed within both studies, separate one-sample Wilcoxon signed-rank tests revealed that respondents in Study 5a chose to blame significantly less than 50% of the time after seeing both clear moral transgressions ( $Median_{\text{choice}} = 0.20, Z = -7.32, p < .001, r = -.49$ ) and ambiguous moral transgressions ( $Median_{\text{choice}} = 0.05, Z = -12.95, p < .001, r = -.88$ ). In Study 6, the rate at which participants chose the BLAME deck did not significantly differ from chance for the clear moral transgressions ( $Median_{\text{choice}} = 0.55, Z = -.28, p = .781, r = -.02$ ), but was significantly less than chance for the ambiguous moral transgressions ( $Median_{\text{choice}} = 0.25, Z = -11.13, p < .001, r = -.77$ ). As revealed by Wilcoxon rank-sum tests, blame choice for clear moral images was significantly higher among respondents in Study 6, who also read captions, than respondents in Study 5a,  $Z = 5.28, p < .001, r = .25$ ; and similar patterns were observed between the two study samples for ambiguous moral images,  $Z = 7.80, p < .001, r = .38$ , which suggests a small-to-moderate effect of captions on blame choice between samples. Overall, even though overall blame preferences were lower in the studies using images, there is a similar pattern within both sets of studies—blame choice tracks the moral ambiguity of the stimuli that is read or seen.

### Cognitive costs

Consistent with results from Studies 5a and 5b, paired sample *t*-tests found that the BLAME deck in Study 6 felt significantly *less* efficacious than the DESCRIBE deck,  $t(209) = -3.23, p = .001, 95\% \text{ CI of } M_{\text{difference}} [-.45, -.11], d = -0.23$ , *more* effortful than the DESCRIBE deck,  $t(209) = 6.69, p < .001, 95\% \text{ CI of } M_{\text{difference}} [.38, .70], d = 0.46$ , and *more* aversive than the DESCRIBE deck,

$t(209) = 6.53, p < .001, 95\%$  CI of  $M_{\text{difference}}$  [.41, .77],  $d = 0.45$ , when images were accompanied by headlines. As the NASA Task Load Index was administered only once in Studies 5a to 6, we did not examine deck differences in cognitive costs by the level of transgression ambiguity.

Across Studies 5a to 6, blaming felt more cognitively challenging in the studies using visual moral stimuli, which is opposite to what was observed in the studies using verbal moral stimuli. In Studies 5a to 6, participants had to interpret potentially unclear visual information, specifically a “snapshot” of an individual in the midst of committing an action. In contrast, the moral vignettes chronologically and explicitly outline a person’s behaviors, thereby lessening the ambiguity and accentuating the harmfulness and moral wrongfulness of their actions. Thus, it follows that, when given the choice, respondents who view the moral images may find it less efficacious, more difficult, and more uncomfortable to blame the individual in the images (and thus, avoid blaming) because it is more challenging to construe a blameworthy action(s). While we acknowledge that images might be more emotionally evocative, the action depicted in the image is likely to be less clear than what is depicted in vignettes. This interpretation is consistent with the socially-regulated blame perspective (Monroe & Malle, 2019), which underscores the role of morally-relevant information about a norm violation (e.g. intentionality) in the systematic and graded up- or down-regulation of blame. Notably, the blame choice effects for clear moral images and for ambiguous moral images were higher in Study 6 than in Study 5a, which suggests that the supplementing verbal information made it easier for respondents to construe moral norm violations from the images and subsequently blame wrongdoers. Together, these results suggest that the willingness to blame appears to track how cognitively costly it feels to do so.

### Study 7: blame selection task with moral vignettes and moral images

The final study extended upon prior studies and assessed relative differences in blame choice and cognitive costs across *both* moral vignettes and moral images. Whereas we observed blame choice differences between different types of moral stimuli across multiple studies, the current study allowed us to compare blame choice between vignettes and images within the *same* group of participants. Building on the pattern of results observed in previous studies, we anticipated that participants’ blame choice would be significantly higher for vignettes than for images. We also expected that blaming would feel significantly more cognitively costly (than the alternative deck) in response to images, and in contrast, significantly less costly (specifically more efficacious

than the alternative deck) in response to vignettes. Between stimuli type, we anticipated that blaming would feel more cognitively costly in response to images than to vignettes. Finally, to confirm that the BST was engendering feelings of outrage, we examined participants’ outrage, arousal, and valence throughout the task to track with their choices to blame.

### Study 7 methods

In Study 7 ( $N = 217$ ), participants saw two blocks of moral stimuli (i.e., images and vignettes) before the card decks (BLAME and DESCRIBE). The order of the blocks were randomized and counterbalanced, such that some participants ( $N = 109$ ) first saw the block of 15 moral images and other participants ( $N = 108$ ) first saw the block of 15 moral vignettes. In this study, we included only the *clear* moral transgressions used in prior vignette and image studies to simplify the study design. For the moral images, we used similar pre-task instructions as Study 5a. For the moral vignettes, we used similar pre-task instructions as Studies 1, 2, and 4a. See the Supplementary Materials for more details.

For the blocks of moral images, participants saw the following trial-level instructions for the two deck choices, BLAME and DESCRIBE: “Try to feel moral outrage toward what you see in the image. Focus on the moral details of the individual and the action they are doing. Please write one sentence to clearly describe why the individual should be blamed” (BLAME); and “Try to be objectively detached toward what you see in the image. Focus on the objective details of the individual and the action they are doing. Please write one sentence to clearly describe what the individual is doing” (DESCRIBE). For the blocks of moral vignettes, participants saw the following trial-level instructions for the two deck choices: “Try to feel moral outrage toward what you read in the scenario. Focus on the moral details of the individual and the action(s) they are doing. Please write one sentence about why the individual should be blamed for this sequence of events” (BLAME); and “Try to be objectively detached toward what you read in the scenario. Focus on the objective details of the individual and the action(s) they are doing. Please write one sentence about how the individual caused this sequence of events” (DESCRIBE). These subtle differences in trial-level instructions between the two types of moral stimuli were meant to retain the instructional sets from previous studies.

After completing the chosen deck’s instructions on each trial, participants then completed outrage, arousal, and valence ratings. We included these ratings to examine whether participants felt more outrage and negative when choosing to blame transgressors, but also to determine whether participants reasonably distinguished between instructions to feel outrage against the alternative.

Participants were asked “How outraged do you feel?” (1 = *not at all*; 9 = *extremely*), “How calm/aroused do you currently feel?” (1 = *calm*; 9 = *aroused*), and “How negative/positive do you currently feel?” (1 = *negative*; 9 = *positive*) using self-assessment manikins (Bradley & Lang, 1994; a similar approach was used to assess state experiences in Scheffer et al., 2021). Further, respondents in Study 7 completed to the NASA Task Load Index twice, once after the first block of trials and once after the second block of trials, in order to collect cognitive costs ratings separately for image-based and vignette-based transgressions.

## Study 7 results

### Blame choice

Collapsing across the counterbalanced conditions, a paired samples Wilcoxon signed rank test found that blame choice for the vignettes was significantly greater than blame choice for the images,  $Z = 10.33$ ,  $p < .001$ ,  $r = .70$ , with a robust effects of stimuli modality observed. See Table 1 for descriptive statistics. Within both counterbalanced conditions, blame choice for the vignettes was significantly greater than blame choice for the images (Images presented first:  $Median_{choice-vignettes} = 0.47$ ,  $Median_{choice-images} = 0.20$ ,  $Z = 6.71$ ,  $p < .001$ ,  $r = .64$ ; Images presented second:  $Median_{choice-vignettes} = 0.57$ ,  $Median_{choice-images} = 0.17$ ,  $Z = 7.87$ ,  $p < .001$ ,  $r = .76$ ), with robust effects of stimuli modality observed. Irrespective of the order of the block of moral images, image-based blame choice was significantly below chance (Images first:  $Z = -7.46$ ,  $p < .001$ ,  $r = -.71$ ; Images second:  $Z = -7.15$ ,  $p < .001$ ,  $r = -.69$ ), which replicates our previous image-based studies (5a to 6) and suggests a strong avoidance among participants to blame transgressors depicted visually. Among participants who saw the block of images first (before the moral vignettes), subsequent blame choice for the vignettes did not significantly differ from chance,  $Z = -1.10$ ,  $p = .272$ ,  $r = -.11$ . Among participants who saw the block of images second (after the moral vignettes), preceding vignette-based blame choice was significantly greater than chance,  $Z = 2.13$ ,  $p = .017$ ,  $r = .20$ , with a small-to-moderate blame choice effect observed.

Taken together, these results are similar to the blame choice patterns observed in prior studies. Across previous studies, we observed that participants approached blame when moral transgressions were depicted verbally, and avoided blame when transgressions were depicted visually. The current study extends upon these patterns and shows that, among the *same* group of participants, the choice to blame is significantly higher for moral vignettes than for moral images. When transgressions are presented in written format, as opposed to visually, participants are more likely to approach outrage and blame. Replicating previous

studies, image-based blame choice was significantly below chance, whereas vignette-based blame choice was significantly above chance.

### Cognitive costs

Table 3 presents the average ratings of perceived effort, perceived aversion, and perceived efficacy for the BLAME and DESCRIBE deck across the image- and vignette-based transgressions in Study 7. As cognitive costs were measured twice, once after each block of moral stimuli, we conducted three  $2 \times 2$  repeated measures ANOVAs to examine whether differences between BLAME and DESCRIBE decks in effort, aversion, and efficacy were moderated by the type of moral stimuli presented to participants. There was a significant interaction between deck (BLAME, DESCRIBE) and stimuli type (images, vignettes) on ratings of efficacy,  $F(1, 216) = 37.67$ ,  $p < .001$ ,  $\eta_p^2 = .15$ , and aversion,  $F(1, 216) = 4.27$ ,  $p = .040$ ,  $\eta_p^2 = .02$ , but not effort,  $F(1, 216) = 2.61$ ,  $p = .108$ ,  $\eta_p^2 = .01$ .

With respect to efficacy, follow-up pairwise comparisons revealed that the BLAME deck felt significantly *more* efficacious than the DESCRIBE deck after the block of vignette-based transgressions ( $M_{difference} = .25$ ,  $SE_{difference} = .08$ ,  $p = .003$ ), and felt significantly *less* efficacious than the DESCRIBE deck after the block of image-based transgressions ( $M_{difference} = -.89$ ,  $SE_{difference} = .09$ ,  $p < .001$ ). Also, the BLAME deck had significantly higher efficacy ratings after the vignette-based trials, compared to after the image-based trials ( $M_{difference} = .65$ ,  $SE_{difference} = .08$ ,  $p < .001$ ), whereas the DESCRIBE deck had significantly higher efficacy ratings after the image-based trials, compared to after the vignette-based trials ( $M_{difference} = .49$ ,  $SE_{difference} = .06$ ,  $p < .001$ ). For aversion, the BLAME deck felt significantly more aversive than the DESCRIBE deck after both the vignettes-block ( $M_{difference} = .42$ ,  $SE_{difference} = .09$ ,  $p < .001$ ) and images-block ( $M_{difference} = .65$ ,  $SE_{difference} = .09$ ,  $p < .001$ ), with the difference between decks being stronger for the image-based

**Table 3** Mean NASA Task Load Index Ratings for BLAME and Alternative decks: Split by Moral Stimuli depicting Transgressions in Study 7

<i>M</i> ( <i>SD</i> ): Study 7	Moral Images	Moral Vignettes
<i>Perceived Effort:</i>		
BLAME deck	3.57 (1.29)	3.51 (1.20)
DESCRIBE deck	3.24 (1.22)	3.35 (1.21)
<i>Perceived Aversion:</i>		
BLAME deck	3.22 (1.40)	3.06 (1.40)
DESCRIBE deck	2.57 (1.29)	2.64 (1.33)
<i>Perceived Efficacy:</i>		
BLAME deck	3.41 (1.28)	4.06 (0.91)
DESCRIBE deck	4.30 (0.73)	3.81 (1.12)



transgressions. The BLAME deck felt marginally more aversive after the images-block than after the vignettes-block ( $M_{\text{difference}} = .16, SE_{\text{difference}} = .08, p = .053$ ); the DESCRIBE deck’s level of aversion did not significantly differ between stimuli type ( $M_{\text{difference}} = -.07, SE_{\text{difference}} = .09, p = .423$ ). Finally, there was a significant main effect for deck on perceived effort,  $F(1, 216) = 11.89, p = .001, \eta_p^2 = .05$ , such that the BLAME deck felt significantly more effortful than the DESCRIBE deck, irrespective of the moral stimuli shown (see Table 3 for mean estimates).

In summary, these findings are consistent with prior vignette-based and image-based studies. The BLAME deck felt significantly *less* costly (more efficacious) than the DESCRIBE deck when participants read the moral vignettes, but felt significantly *more* cognitive costly (more effortful, aversive, and inefficacious) than the DESCRIBE deck when participants saw the moral images. Moreover, between both types of moral stimuli, the BLAME felt less costly to use (specifically, more efficacious and less aversive)

**Table 4** Multilevel Models Predicting Changes in Outrage, Arousal, and Valence from Choice and Stimuli Type: Study 7

Outcome	Effect	<i>F</i>	<i>df</i>	<i>p</i>
<b>Outrage:</b>	Choice	1248.73	6410.45	<.001
	Stimuli Type	7.59	6305.84	.006
	Choice * Stimuli Type	19.98	6336.11	<.001
<b>Arousal:</b>	Choice	755.59	6395.17	<.001
	Stimuli Type	131.91	6303.62	<.001
	Choice * Stimuli Type	.25	6329.50	.615
<b>Valence:</b>	Choice	319.36	6392.94	<.001
	Stimuli Type	60.50	6303.15	<.001
	Choice * Stimuli Type	4.63	6328.46	.032

The table presents Type III Tests of Fixed Effects

when people saw vignettes (as opposed to images), which suggests a greater felt cost to blaming transgressors depicted in images.

**Outrage and state affect ratings**

Finally, we examined the degree to which choosing blame influenced ratings of outrage, arousal, and valence, additionally modeling the interaction between deck choice and moral stimuli type (image, vignette). For this analysis, we conducted multilevel models using SPSS MIXED, nesting trials within participants. Each model (*N* observations = 6510) predicted changes in state ratings (outrage, arousal or valence) from choice (describe = 0, blame = 1), type of moral stimuli (vignettes = - 1, images = 1), and the interactive effect of choice and type of moral stimuli. Each model included the fixed effects of intercept, choice, stimuli type, and the interaction term, as well as a random intercept. In this main analysis, we did not model repeated measure of time because of the blocking of the stimuli type manipulation (confounding stimuli condition with time).<sup>3</sup> Details of these analyses are presented in Table 4 (Type III Tests of Fixed Effects) and Table 5 (Estimated Marginal Means).

The estimates of fixed effects noted below are based on when the choice = 0 (i.e., when the DESCRIBE deck is selected) and when stimuli type = - 1 (i.e., when vignettes are shown). With the contrast between blame and objective detachment, there were significant effects of deck choice,  $B = -2.23, SE = .08, t = -27.31, p < .001, 95\% \text{ CI } [-2.39, -2.07]$ , moral stimuli type,  $B = -.37, SE = .08, t = -4.53, p < .001, 95\% \text{ CI } [-.53, -.21]$ , and the choice × stimuli type interaction,  $B = .46, SE = .10, t = 4.47, p < .001, 95\% \text{ CI } [.26, .67]$ , on changes in outrage. Similarly, there were significant effects of deck choice,  $B = .85, SE = .06, t = 13.75, p < .001, 95\% \text{ CI } [.73, .98]$ , moral stimuli

**Table 5** Estimated Marginal Means from Multilevel Models Predicting Changes in Outrage, Arousal, and Valence from Choice and Stimuli Type: Study 7

Outcome	Stimuli Type	BLAME deck: <i>M</i> ( <i>SE</i> )	DESCRIBE deck: <i>M</i> ( <i>SE</i> )	<i>p</i>	95% CI <i>M<sub>diff</sub></i>
<b>Outrage:</b>	Vignette	6.24 (.13)	4.47 (.13)	<.001	[1.63, 1.91]
	Image	6.61 (.14)	4.38 (.13)	<.001	[2.07, 2.39]
<b>Arousal:</b>	Vignette	5.07 (.13)	3.62 (.13)	<.001	[1.32, 1.59]
	Image	5.65 (.14)	4.14 (.13)	<.001	[1.35, 1.66]
<b>Valence:</b>	Vignette	2.81 (.11)	3.50 (.11)	<.001	[-0.79, -0.58]
	Image	2.43 (.11)	3.29 (.11)	<.001	[-0.98, -0.73]

<sup>3</sup> In secondary analyses, which modelled time as a repeated measure with an autoregressive covariance structure (time was coded to reflect trials 1–30, such that trials 1–15 reflect image-based transgressions for some participants and vignette-based transgressions for other participants, and vice versa for trials 16–30), results for the choice × stimuli type interactions were similar for the fixed effects and pairwise comparisons between estimated marginal means.

**Table 6** Meta-analytic Correlations between Blame Choice, Cognitive Costs, Norms and Individual Differences

Measure	Blame Choice	<i>p</i>	Effort	<i>p</i>	Aversion	<i>p</i>	Efficacy	<i>p</i>
Blame Choice	–		–.12 [–.17, –.06]	<.001	–.14 [–.19, –.09]	<.001	.50 [.43, .56]	<.001
Descriptive Norm	.46 [.40, .51]	<.001	–.17 [–.23, –.11]	<.001	–.14 [–.19, –.09]	<.001	.33 [.26, .39]	<.001
Injunctive Norm	.10 [.05, .15]	<.001	–.02 [–.08, .05]	.622	–.02 [–.07, .03]	.348	.09 [.04, .14]	<.001
Personal Desire	.29 [.21, .36]	<.001	–.05 [–.11, .02]	.153	–.11 [–.16, .06]	<.001	.23 [.18, .28]	<.001
Moral Convictions	.30 [.20, .40]	<.001	–.04 [–.09, .02]	.157	–.04 [–.09, .01]	.119	.17 [.10, .24]	<.001
Symbolization	.15 [.09, .22]	<.001	–.03 [–.09, .03]	.327	–.04 [–.10, .01]	.132	.14 [.09, .19]	<.001
Internalization	.03 [–.04, .09]	.411	–.03 [–.07, .02]	.323	.00 [–.05, .05]	.973	–.04 [–.10, .03]	.322
Empathy	.12 [.04, .20]	.005	–.02 [–.09, .04]	.456	–.04 [–.10, .03]	.242	.09 [.03, .15]	.003
Beh. Contagion	.09 [.02, .15]	.009	.01 [–.05, .07]	.775	–.01 [–.04, .06]	.675	.07 [.02, .13]	.005

Meta-analytic effects reflect Pearson's *r* correlations, with 95% confidence intervals presented. Perceived Effort, Aversion, and Efficacy, as well as Descriptive and Injunctive norms, reflect difference scores (Blame deck – Alternative deck). For Blame Choice correlations involving NASA ratings of Effort, Aversion, and Efficacy, Moral Convictions, Empathy, Behavioral Contagion, *N*s = 1672. For Blame Choice correlations involving Symbolization (moral identity subscale), Internalization (moral identity subscale), Descriptive Norm, Injunctive Norm, and Personal Desire, *N*s = 1570. In Studies 3 and 7, blame choice and cognitive cost measures are averaged across the within-subjects transgression ambiguity condition (Study 3: clear, ambiguous) and stimuli type condition (Study 7: images, vignettes)

type,  $B = .38$ ,  $SE = .06$ ,  $t = 6.13$ ,  $p < .001$ , 95% CI [.26, .50], and the choice  $\times$  stimuli type interaction,  $B = -.17$ ,  $SE = .08$ ,  $t = -2.15$ ,  $p = .032$ , 95% CI [–.32, –.02], on changes in valence. Participants reported lower levels of outrage and increased positive valence after choosing the DESCRIBE deck (than after choosing the BLAME deck) and these differences were amplified on trials involving moral vignettes (compared to moral images). Put differently, participants felt *more* outrage and negative valence after choosing the BLAME deck (versus the DESCRIBE deck) and these differences were amplified on trials involving *moral images* (compared to moral vignettes). For arousal, there were significant effects for choice,  $B = -1.50$ ,  $SE = .08$ ,  $t = -19.36$ ,  $p < .001$ , 95% CI [–1.66, –1.35], and stimuli type,  $B = -.57$ ,  $SE = .08$ ,  $t = -7.35$ ,  $p < .001$ , 95% CI [–.72, –.42], but not for the choice  $\times$  stimuli type interaction,  $B = .05$ ,  $SE = .10$ ,  $t = .50$ ,  $p = .615$ , 95% CI [–.14, .24]. Likewise, participants reported *higher* levels of arousal after choosing the BLAME deck (than after choosing the DESCRIBE deck) and they reported *higher* levels of arousal after seeing the moral images (than after reading the moral vignettes). In summary, participants who chose blame felt more outraged, more negative, and more aroused (as opposed to calm), providing further evidence that blame regulation had consequences for generating outrage. These results also provide a manipulation check: Choosing to blame (instead of remaining objectively detached) led participants to feel more outrage towards the transgressor.

### Meta-analytic associations of blame choice with cognitive costs, norms, and individual differences

In this section, we explored the relationships between blame choice and perceived cognitive costs measured with the NASA Task Load Index, which allowed us to test the hypothesis that

to the degree that participants felt blaming (relative to the alternative deck) was cognitively costly, they would not choose to blame wrongdoers. Although overall blame choice patterns differed between the studies using vignettes and the studies using images, it does not preclude the possibility that blame choice would have consistent associations with other constructs across studies, such as people's felt cognitive costs. In order to garner more statistical power, we computed meta-analytic correlations across all studies of blame choice with: perceived cognitive costs measured by the NASA Task Load Index (i.e., perceived effort, aversion, and efficacy assigned to the BLAME deck versus the alternative deck); and with descriptive, injunctive, and personal norms about blaming (versus the alternative deck). We also computed the meta-analytic correlations of blame choice with stable individual differences measured across the studies, such as moral convictions (Moral Convictions Scale; Skitka, 2010), moral identity (Moral Identity Scale; Aquino & Reed, 2002), and empathy (Empathy Index; Jordan et al., 2016a).

For the NASA cognitive costs indices and descriptive and injunctive norms, difference scores between the BLAME deck and the alternative deck were computed (e.g., “BLAME deck Perceived Effort” minus “Alternative deck Perceived Effort”). The Hedges–Olkin method (Borenstein et al., 2009) calculated an overall correlation coefficient (*r*) from the set of correlations across all studies (using Comprehensive Meta-Analysis; Borenstein et al., 2013). Table 6 summarizes the meta-analytic effects,

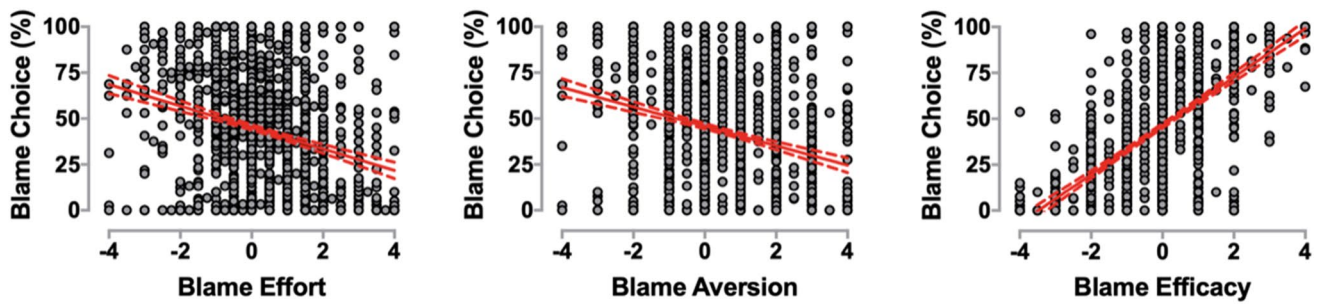


Fig. 5 Associations of Blame Choice with NASA Task Load Index Ratings of Perceived Effort, Aversion, and Efficacy of Blaming (versus Objectivity or Empathy)

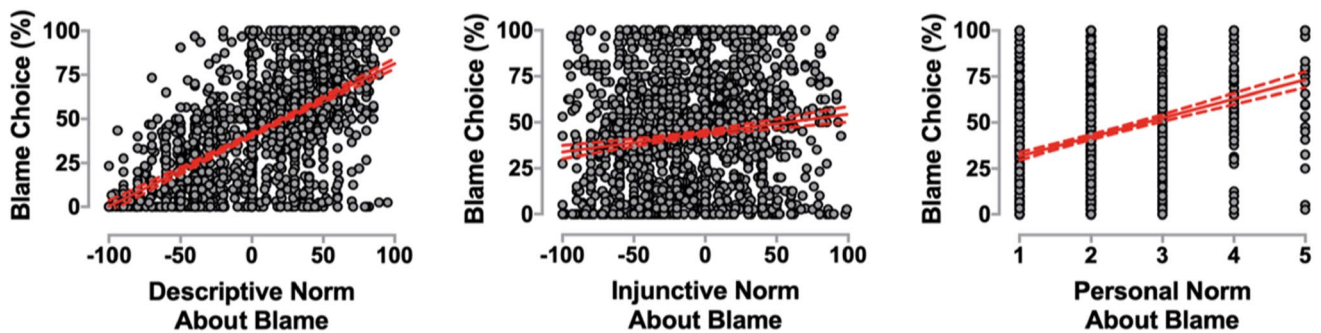


Fig. 6 Associations of Blame Choice with Descriptive, Injunctive, and Personal Norms about Blaming (versus Objectivity or Empathy)

including the 95% confidence intervals. Table S9 in the Supplementary Materials presents the within-study correlations of blame choice with all of the above-noted measures.

First, with respect to cognitive costs,<sup>4</sup> blame choice was negatively related to perceived effort,  $Z = -3.85$ ,  $r = -.12$ ,  $p < .001$ , 95% CI  $[-.17, -.06]$ , and perceived aversion,  $Z = -5.27$ ,  $r = -.14$ ,  $p < .001$ , 95% CI  $[-.19, -.09]$ . Blame choice was positively related to perceived efficacy,  $Z = 12.64$ ,  $r = .50$ ,  $p < .001$ , 95% CI  $[.43, .56]$ . In other words, respondents were less likely to choose the BLAME deck when they viewed the deck as more effortful and aversive, as well as less efficacious, than the DESCRIBE or FEEL deck. Figure 5

displays each of these associations aggregated across all studies. Tests of the difference between two dependent correlations (with one variable in common; Steiger, 1980) revealed that the association of blame choice with perceived efficacy was significantly more robust than the associations of blame choice with perceived effort,  $Z = 17.51$ ,  $p < .001$ , and perceived aversion,  $Z = 18.08$ ,  $p < .001$ . Taken together, these results provide further support for our prediction that the cognitive costs of blaming—especially felt efficacy—are related with respondents’ choices to blame individuals. One’s choice to feel outrage and blame transgressors is linked to the degree that doing so feels less effortful, less aversive, and more efficacious than being objective or empathizing.

<sup>4</sup> Given the different patterns of blame choice between the vignette and image studies, we also report separate meta-analytic correlations between blame choice and cognitive costs for the vignette and images studies, as the NASA Task Load Index items were rated based on the BST that respondents completed prior. The following patterns were observed within each study set: *Vignette studies* (blame choice with perceived effort:  $r = -.13$ ,  $p < .001$ ; blame choice with perceived aversion:  $r = -.19$ ,  $p < .001$ ; blame choice with perceived efficacy:  $r = .55$ ,  $p < .001$ ) and *Image studies* (blame choice with effort:  $r = -.05$ ,  $p = .480$ ; blame choice with aversion:  $r = -.05$ ,  $p = .151$ ; blame choice with efficacy:  $r = .41$ ,  $p < .001$ ).

Second, blame choice was positively associated with all social and personal blaming norms<sup>5</sup> i.e., the descriptive norm difference score,  $Z = 13.67$ ,  $r = .46$ ,  $p < .001$ , 95% CI [.40, .51], the injunctive norm difference score,  $Z = 4.08$ ,  $r = .10$ ,  $p < .001$ , 95% CI [.05, .15], and the personal belief that blaming is more desirable than the alternative choice,  $Z = 7.36$ ,  $r = .29$ ,  $p < .001$ , 95% CI [.21, .36]. That is, to the degree that respondents believed that other participants choose the BLAME deck, as opposed to the DESCRIBE or FEEL deck, there was a moderate-to-large increase in respondents' own blame choice. To the extent that other participants believe that blaming, as opposed to remaining objective or empathizing, is a "good thing", there was a small increase in respondents' own blame choice. Finally, to the degree that respondents personally believed that blaming was more desirable than being objective or empathic, there was a moderate increase in their blame choice. Figure 6 displays each of these associations aggregated across all studies. These correlations of blame choice with descriptive and injunctive social norms are consistent with results that highlight the signalling function of expressing moral outrage (Jordan et al., 2016b), as well as results that demonstrate the link between group preferences to punish and respondents' own punitive decisions when tasked with restoring justice (Son et al., 2019). Here, the choice to blame appears to track the social value of doing so. Given that these are only correlational results, it is possible that social and personal norm ratings reflect respondents' post-hoc justifications of blame choices made earlier in the task.

Lastly, blame choice was positively associated with moral convictions,  $Z = 5.61$ ,  $r = .30$ ,  $p < .001$ , 95% CI [.20, .40], and symbolized moral identity,  $Z = 4.84$ ,  $r = .15$ ,  $p < .001$ , 95% CI [.09, .22], but not with internalized moral identity,  $Z = .82$ ,  $r = .03$ ,  $p = .411$ , 95% CI [−.04, .09]. In other words, participants who reported having stronger moral convictions engaged by the task demonstrated a moderate increase in their choice to blame; and among participants who reported engaging in actions that socially and symbolically demonstrate their moral identity (i.e., behaving in ways that reflect moral traits), there was a modest increase in blame choice. Blame choice was also positively, though weakly, associated with both subscales of the Empathy Index: Empathy,

$Z = 2.84$ ,  $r = .12$ ,  $p = .005$ , 95% CI [.04, .20], and Behavioral Contagion,  $Z = 2.60$ ,  $r = .09$ ,  $p = .009$ , 95% CI [.02, .15]. These results suggest that participants more prone to taking on others' feelings and behaviors have a modest increase in the tendency to blame perpetrators of transgressions. Finally, blame choice was positively associated with the Blame Intensity Inventory (Gill & Cerce, 2021),  $r = .21$ ,  $p = .002$  (included only in Study 7), which suggests that participants with a propensity to experience hostile affect and reactions towards wrongdoers demonstrated a small-to-moderate increase in their choice to blame wrongdoers. The correlations of blame choice with blaming norms and with multiple indices of moral personality provide evidence for the convergent validity for the BST.

Taken as a whole, these results suggest that respondents' felt cognitive costs, their prescribed norms about blaming, as well as their moral personality and empathic traits were related to their blame choice. The choice to blame the individuals depicted in the moral stimuli was most strongly linked to the perceived efficacy of the BLAME deck and to the belief that other participants would choose the BLAME deck (as opposed to the alternative option).

## Discussion

Moral norm violations cause people to experience moral outrage and to express it in various ways (Crockett, 2017), such as shaming/dehumanizing, punishing, or blaming. These forms of expressing outrage are less than moderately related to one another ( $r$ 's  $< .30$ ; see Bastian et al., 2013 for more information), which suggests that a considerable amount of variance between shaming/dehumanizing, punishing, and blaming remains unexplained and that these are distinct enough demonstrations of outrage in response to norm violations. Yet, despite its moralistic implications (see Crockett, 2017), there is still little empirical work not only on the phenomenon of outrage fatigue but also on the role of motivated cognition on expressing outrage via blame. Social costs alter blame judgments, even when people's cognitive resources are depleted (Monroe & Malle, 2019). But how do the inherent cognitive costs of blaming relate to people's decisions towards moral outrage and blame? Here, we examined how felt cognitive costs associate with the choice to express outrage through blame.

Taking our cues from the Law of Least Effort (Hull, 1943; Kool et al., 2010), we developed a free-choice measure of blame-proneness—the BST—to observe how motivated respondents were to generate outrage and blame individuals committing moral transgressions. This task allows people to select whether they want to enter into situations (i.e., situation selection; Gross & Thompson, 2007) that call for outrage cultivation and blaming, and it has parallels to everyday

<sup>5</sup> While we have presented the overall meta-analytic correlations, we also report separate meta-analytic correlations between blame choice and social/personal norms for the vignette and images studies, as the norm items were rated based on the BST that respondents completed prior. The following patterns were observed within each study set: *Vignette studies* (blame choice with descriptive norms:  $r = .48$ ,  $p < .001$ ; blame choice with injunctive norms:  $r = .09$ ,  $p = .006$ ; blame choice with personal norms:  $r = .29$ ,  $p < .001$ ) and *Image studies* (blame choice with descriptive norms:  $r = .39$ ,  $p < .001$ ; blame choice with injunctive norms:  $r = .11$ ,  $p = .001$ ; blame choice with personal norms:  $r = .27$ ,  $p < .001$ ).

examples of either avoiding news about tragedies and harms or consciously “doom-scrolling” to up-regulate outrage (frequently observed on Twitter; see Miller, 2020). The BST relies on behavioral measurement of outrage and blame as people choose to regulate it, rather than participants’ self-reports. This free-choice approach provides a novel method for understanding motivated regulation of blame and its associated cognitive costs.

With respect to blame choice patterns, participants preferred to blame at a rate greater than 50% in the studies that used moral vignettes. Most notably, when the decks preceded the vignettes, respondents displayed a significant small-to-moderate preference to choose the BLAME deck—irrespective of if the alternative choice was to objectively write about how the individual caused the sequence of events or to empathically write about the individual’s feelings and thoughts. In other words, participants made the choice to express outrage through blame, as opposed to remain objective or to feel empathy, even before reading the moral vignette and the individual’s behavior(s). Irrespective of the alternative deck, participants also displayed a significant blame preference when the vignettes preceded the decks. Particularly, they displayed a significantly higher preference to choose the BLAME deck when the vignettes were blatantly harmful and immoral (clear moral transgressions) than when the moral transgressions were ambiguous, with a very large effect of moral ambiguity on blame choice observed. Whereas participants blamed significantly more than 50% of the time for the clear moral violations, with a large blame choice effect observed, they did not show a preference for blaming individuals of ambiguous moral violations, with a very small blame choice effect observed. These findings extend upon those by Jordan and Kteily (2020) by examining people’s personal regulation of outrage—outside of the perceived reputational incentives of outrage—and they run parallel with literature that links observers’ perceived harm of wrongdoings to the severity of their retributive punishment of transgressors (e.g., Darley & Pittman, 2003).

In contrast, within the studies that used moral images, individuals chose to blame individuals at a rate significantly less than 50%—irrespective of if the decks preceded or followed the images; here, a moderate-to-very large blame avoidance effect was observed. In studies where the decks followed the moral images (Studies 5a and 6), blame selection was significantly greater for clear moral transgressions—replicating results from the moral vignette studies—with a very large effect of moral ambiguity on blame choice. Yet, in contrast to the vignette studies, blame choice for both clear and ambiguous moral images did not significantly exceed chance. Together, these blame choice findings suggest that the change in moral stimuli, from vignettes to images, contributed to participants’ avoidance of blaming the people depicted in the stimuli. Indeed, Study 7 found

that blame choice for vignette-based moral transgressions was significantly greater than blame choice for image-based moral transgressors among the same group of participants, suggesting a motivation towards blaming when actions are depicted verbally, as opposed to visually. Of note, blame choice for the vignette-based trials exceeded chance when these trials preceded the image-based trials, but did not exceed chance when these trials were preceded by the image-based trials. This pattern could potentially be accounted for by a “carryover effect” of blame avoidance. That is, among those who first saw the block of moral images, it is possible that these participants were less motivated to subsequently blame for the moral vignettes because they had developed a pattern of low blame choice for the images.

Notably, these studies are among the first to examine whether generating outrage and blame feels cognitively costly—without any external manipulations of effort (Ames & Fiske, 2015). Our results are the first to examine whether and to what degree cognitive costs relate to this moralizing process. The consistent finding across the studies that used moral vignettes was that blaming felt more efficacious than the alternative choice, which could account for why participants’ blame choice exceeded chance in these studies. Across the studies that used moral images, blaming felt less efficacious and more effortful and aversive than the alternative choice, which could account for why participants’ blame choice did not exceed chance in these studies. In Study 7, blaming felt more efficacious and less aversive after the vignette-based trials than after the image-based trials, which implies that it felt more costly to generate outrage and blame when transgressions were depicted visually (versus visually). The meta-analytic associations found that people were less likely to choose the BLAME deck when it felt more cognitively effortful and aversive than the alternative deck, and more likely to choose BLAME when it felt more efficacious. Our results are the first to suggest that blaming relates to how cognitively costly it feels to do so.

Of the three facets of cognitive costs measured by the NASA Task Load Index, perceived efficacy was the strongest correlate of blame choice—across all studies and within each study set (vignettes and images). Indeed, squaring each correlation suggests that respondents’ perceived efficacy of the BLAME deck accounted for 16.81-to-30.25% of the variance in their choice to blame, compared to 0.25-to-1.69% and 0.25-to-3.61% of the variance from perceived effort and perceived aversion, respectively. This indicates that one’s perceived inefficacy of blaming (versus remaining objective or empathizing) is a more critical aspect of cognitive work in the choice to express outrage and blame than one’s perceived effort and aversion. One might desire to avoid inaccurate moral decisions due to the social costs (Jordan et al., 2016b; Kundu & Cummins, 2013), but our results highlight the link between one’s choice to blame and how inherently

successful or worthwhile blaming feels. This finding is likely to have important moralistic implications, as people's perceptions of efficacy have been found to bolster their commitment to collective pro-social actions within moral contexts (Thomas & McGarty, 2009).

Across studies and regardless of stimulus modality, participants were more likely to choose to blame for clear moral transgressions than for ambiguous transgressions, and the differences in felt cognitive costs across transgression type noted in Study 3 might suggest one reason why. For clear moral transgressions in Study 3, participants rated blaming as significantly more efficacious and less effortful. For ambiguous moral transgressions, where there might be conflicting ideas about how to respond, the cognitive ease of blame reduced. Moreover, when faced with more equivocal stimuli in general—the visual depictions of transgressions—participants avoided blaming instead of choosing it. The cognitive cost differences between decks observed in the studies that used images, as well as the differences between the vignette- and image-based trials in blame-related efficacy and aversion found in Study 7, suggest that the uncertainty involved with interpreting visual stimuli made blaming more cognitively demanding. The socially-regulated blame perspective posits that blame must be warranted by morally-relevant information in order for individuals to circumvent the social costs of blaming (Monroe & Malle, 2019), and this is relevant in light of the meta-analytic correlations between social norms and blame choice. Across all studies, participants' responses were anonymous and their choices were not shared with others. As such, the blame choice patterns between clear versus ambiguous moral transgressions and between the studies that used vignettes versus images suggest that, even in the absence of social costs, people are sensitive to whether or not there is sufficient warrant to blame wrongdoers. This, in turn, could account for the higher cognitive costs attached to blame in the image-based studies and the larger blame avoidance in these studies.

Although much work has explored how external factors influence blame judgements (i.e., transgressor causality, intentionality, preventability, harm, other circumstances, etc.; Malle et al., 2014), very little work has examined individuals' baseline preferences to blame and how such preferences track features of the context (i.e., stimuli type, transgression ambiguity, and alternative deck) in interaction with individual characteristics. Supporting the claim that the BST captures the motivation to moralize, participants more deeply convicted in their moral principles and beliefs about fundamental right and wrong (moral convictions) showed a greater tendency to choose to blame across studies. Relatedly, Study 7 found that those prone to experiencing hostile affect towards offenders (blame intensity) chose to blame more. Moreover, blame choice appeared sensitive to social aspects of moralizing, with blame selection associating with

individuals' desire to act in accordance with moral traits for its social and public consequences (symbolic moral identity). Further, participants more prone to take on the feelings (empathy) and manners (behavioral contagion) of others demonstrated higher blame preference, which suggests that, in the context of expressing outrage, empathic individuals are choosing to empathize with the victims of transgressions, not the perpetrators. As noted above, blame choice appeared sensitive to individuals' descriptive and injunctive norms about whether other people choose to blame and value blame over remaining objective or empathizing, as well personal beliefs about the appeal of blame. These results not only cohere with other results that suggest a strong signalling function for expressing outrage (e.g., Jordan et al., 2016b) and a strong link between group preferences for expressing outrage and one's own decisions (Son et al., 2019), but also highlight another pathway through which people commit to collective moral action—group norms (Thomas & McGarty, 2009). Taken as a whole, the associations of blame choice with all of these individual difference measures provide support for the convergent validity of the BST.

When modelling the interactive effect of choice (blame, describe) and stimuli type (vignettes and images) on changes in outrage, arousal, and valence in Study 7, the multilevel models found that participants reported higher outrage and negative valence after choosing to blame (than describe), and that these differences were most pronounced on the trials that used moral images. Additionally, participants' arousal levels were higher after choosing to blame (as opposed to describe) and after seeing moral images (than after reading moral vignettes). The effects of choice on increases in outrage, arousal, and negative valence provides support for the construct validity of the BST: Participants had different levels of outrage, arousal, and valence when choosing blame instead of objectivity, which suggests that the directions participants were provided, when chosen, led to different outrage responses. Interestingly, even though images provoked *increases* in outrage, negative valence, and arousal, participants chose to blame at a rate significantly *less* than chance for the image-based transgressions. We posit that the blame choice for images is lower than chance because the action depicted in the image may be more difficult to construe and hence, more difficult to blame. Given the links between affective intensity and people's level of cognitive flexibility and processing (e.g., Hsieh & Lin, 2019; Vásquez-Rosati et al., 2019), it is additionally plausible that these engendered negative emotions (i.e., outrage) hindered respondents' ability to construe the wrongdoing and/or generate valid reasons as to why the person depicted in the image should be blamed for their actions, which, in turn, contributed to a stronger blame avoidance among respondents.

## Limitations and future directions

We acknowledge that task demand or felt cognitive costs may not deter blaming if individuals' motivations to blame have not been satisfied. For example, intentional harm-doings elicit the motivation to blame, condemn, and punish the harm-doers, and motivated participants will initially raise their level of effort to match the increasing task demands to express outrage (Ames & Fiske, 2015). Further, whether or not people find blaming to be demanding will likely vary depending on the context. We varied the contrast deck (i.e., objectivity or empathy) and whether individuals were moralizing specific actions (i.e., moral stimuli before choice) or moralizing in general (i.e., moral stimuli after choice), and these did not make a substantial impact on blame choice. Nevertheless, there could be other contextual factors, such as the amount of causal and mental information provided about the transgressions, that matter for moral decision-making and associated felt cognitive costs. Moreover, some work finds that costly punishment can signal trust to others, but that this changes when costly helping is provided as an alternative behavioral option (Jordan et al., 2016b). This finding suggests that individuals may be sensitive to the signalling value of outrage expression (versus empathy) in social interactions, and future studies might productively explore how people anticipate the social impact of effort-based choices to select or avoid blame. Importantly, we acknowledge that the results of the current study do not generalize to all demonstrations of moral outrage. Thus, future iterations of the task will want to explore participants' baseline motivations to shame/dehumanize or to punish transgressors, the felt cognitive costs associated with these different ways of expressing outrage, and if such costs differ from those observed with blaming.

In the current studies, the repeated-measures assessment of blame choice during the BST precluded us from examining the cognitive costs of each deck until *after* the task was completed and all blame choices had been made. As such, we acknowledge that this leaves open the possibility that cognitive dissonance led those participants who frequently chose to blame to perceive blaming as less cognitively costly (than the alternative deck) after the task, than those participants who did not blame as frequently. That is, participants' post-task rationalization for choosing to blame could account for the relationship between blame choice and ratings of cognitive costs. Exploratory analyses in Study 3 suggest that, among respondents who first read the block of ambiguous moral transgressions, their blame choice during the second half of the BST was linked to the degree to which blaming felt less aversive ( $r = -.27, p = .007$ ), less effortful ( $r = -.30, p = .001$ ), and more efficacious ( $r = .32, p = .001$ ) than empathy during the first half of the BST (while statistically controlling for cognitive cost differences during the second half of the task). A similar pattern was observed

between felt efficacy and subsequent blame choice ( $r = .25, p = .013$ ) among respondents who first read the clear transgressions. While these exploratory results suggest that people avoid blaming when it feels more cognitively costly, they also underscore the need to further explore how blame choice and felt cognitive costs causally influence each other. As recent work finds that experimentally increasing felt efficacy eliminates empathy avoidance (Cameron et al., 2019), it is possible that manipulating the efficacy of blaming might similarly shape choices to blame. Future work should utilize procedures to experimentally increase or decrease the felt cognitive costs of blaming and examine how this subsequently impacts blame selection.

Finally, we acknowledge that the BST, in its current iterations, may have low external validity and applicability to everyday contexts of outrage, such as the kinds of outrage observed on social media (e.g., posting, reposting, 'liking', or 'reacting' to content). Although the current variations of the BST may lack some degree of external validity, there are trade-offs between external and internal validity (see Aronson et al., 1998), and we argue that this concern of external validity may still be tolerable if a research method can provide an internally valid test of the research questions of interest (i.e., by providing a structured assessment of how individuals chose to engage with outrage/blaming or not). In other words, there can, at times, be advantages of a context-limited method if such method provides a useful and valid test of psychological phenomena. Also, the repeated measurement of choices within our task builds from previous work that uses repeated trials to establish reliable estimates of behavioral preferences (e.g., Kool et al., 2010). Along with the realistic examples of moral transgressions depicted verbally and visually, our task is representative of the types of moral stimuli (e.g., images, brief verbal descriptions) and moral wrongdoings (e.g., racism, infidelity) that people encounter on online platforms, as well as the frequency of the encounters that people have with such moral content on these platforms (e.g., Crockett, 2017; Hofmann et al., 2014).

Although we did not examine real-world outrage behaviors (e.g., retweeting), it is promising that our work complements emerging work on real-world behaviors on social media. For instance, studies reveal that Twitter posts with moral outrage (Brady et al., 2020a, b) or hostility towards perceived wrongdoers (e.g., political opponents; Rathje et al., 2021) are more frequently distributed online and that people choose to associate with others who express more extreme levels of outrage online (Goldenberg et al., 2020). To further establish the validity of the BST, future studies will want to examine the link between individuals' blame choice behaviors during the task and their outrage-related behaviors in real-world contexts.

## Conclusions

Are feeling outrage and blaming effortless and intuitive, or effortful and fatiguing? When confronted with written accounts of clear moral transgressions, individuals show a preference to blame and they feel more efficacious and skilled at doing so, in comparison to objectively or empathically engaging with written indiscretions. When confronted with visual depictions of moral transgressions, individuals avoid blaming and they feel less efficacious and skilled at doing so, as well as more cognitively taxed and negative. It feels cognitive easier to blame when moral wrongdoings are depicted verbally, as opposed to visually, and thus, when individuals encounter both types of moral content, they show a blame choice preference towards verbal information. Yet, our studies are the first to demonstrate that even as the choice to blame and its felt cognitive costs vary across contexts, these two experiences track together clearly: Individuals choose to blame more often when it feels less effortful, less aversive, and particularly more efficacious. Outrage and blame may seem easy or intuitive to engage in, given the platforms and influx of offensive and immoral stories that are accessible to individuals. Our research advances the study of blame by suggesting that individuals choose to feel outrage and blame when it feels like cognitive ease, rather than cognitive work.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11031-021-09917-y>.

**Author contributions** All authors contributed to the conception and design of studies. Material preparation, data collection, and data analyses were performed by VB. Drafts and revisions of the manuscript were completed by VB and all authors provided feedback and commented on the manuscript. All authors read and approved the final manuscript.

**Funding** This research was supported by Grant 1660707 from the National Science Foundation and Grant 61150 from the John Templeton Foundation to the 2nd author, and Grant 435-2014-0556 from the Social Sciences and Humanities Research Council and Grant RGPIN-2014-03744 from the Natural Sciences and Engineering Research Council to the 3rd author.

## Declarations

**Conflict of interest** The authors have no significant interests to report.

## References

- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126(4), 556–574. <https://doi.org/10.1037/0033-2909.126.4.556>
- Alicke, M. D. (2008). Blaming badly. *Journal of Cognition and Culture*, 8(1), 179–186. <https://doi.org/10.1163/156770908x289279>
- Ames, D. L., & Fiske, S. T. (2015). Perceived intent motivates people to magnify observed harms. *Proceedings of the National Academy of Sciences of USA*, 112(12), 3599–3605. <https://doi.org/10.1073/pnas.1501592112>
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037//0022-3514.83.6.1423>
- Aronson, E., Wilson, T. D., & Brewer, M. B. (1998). Experimentation in social psychology. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 99–142). McGraw-Hill.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 1531–1542. <https://doi.org/10.1177/0956797615594620>
- Bastian, B., Denson, T. F., & Haslam, N. (2013). The roles of dehumanization and moral outrage in retributive justice. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0061842>
- Bastian, B., Laham, S. M., Wilson, S., Haslam, N., & Koval, P. (2011). Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology*, 50(3), 469–483. <https://doi.org/10.1348/014466610x521383>
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2013). *Comprehensive Meta-Analysis Version 3*. Biostat.
- Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020a). The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspectives on Psychological Science*, 15(4), 978–1010. <https://doi.org/10.1177/1745691620917336>
- Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020b). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, 149(4), 746–756. <https://doi.org/10.1037/xge0000673>
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences of USA*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>
- Cameron, C. D., Hutcherson, C. A., Ferguson, A. M., Scheffer, J. A., Hadjiandreou, E., & Inzlicht, M. (2019). Empathy is hard work: People choose to avoid empathy because of its cognitive costs. *Journal of Experimental Psychology: General*, 148(6), 962–976. <https://doi.org/10.1037/xge0000595>
- 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(2), 284–299. <https://doi.org/10.1037//0022-3514.83.2.284>
- Crockett, M. J. (2017). Moral outrage in the digital age. *Nature Human Behavior*, 1(11), 769–771. <https://doi.org/10.1038/s41562-017-0213-3>
- Crone, D. L., Bode, S., Murawski, C., & Laham, S. (2018). The Socio-Moral Image database: A novel stimulus set for the study of social, moral and affective processes. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0190954>
- Darley, J. M., & Pittman, T. S. (2003). The psychology of compensatory and retributive justice. *Personality and Social Psychology Review*, 7(4), 324–336. [https://doi.org/10.1207/S15327957PSPR0704\\_05](https://doi.org/10.1207/S15327957PSPR0704_05)
- Diamond, J. (2017, January 29). Trump's latest executive order: Banning people from 7 countries and more. *CNN*. <https://www.cnn.com/2017/01/27/politics/donald-trump-refugees-executive-order/index.html>
- Gardner, T., & Monroe, A. E. (2018). Blaming the blamers: People impose different obligations to punish on victims versus third



- parties. In *Poster presented at the Society of Southeastern Social Psychologists conference*, Raleigh, NC.
- Gill, M. J., & Cerce, S. C. (2021). The Blame Intensity Inventory: Assessing the propensity to blame harshly and its unique capacity to predict malicious satisfaction from offender victimization. *Personality and Social Psychology Bulletin*, 47(12), 1668–1685. <https://doi.org/10.1177/0146167220985362>
- Gill, M. J., & Getty, P. D. (2016). On shifting the blame to humanity: Historicist narratives regarding transgressors evoke compassion for the transgressor but disdain for humanity. *British Journal of Social Psychology*, 55(4), 773–791. <https://doi.org/10.1111/bjso.12159>
- Goldenberg, A., Abruzzo, J., Willer, R., Halperin, E., & Gross, J. (2020). Homophily and acrophily as drivers of political segregation. <https://doi.org/10.31219/osf.io/9xetr>
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)
- Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3–24). Guilford Press.
- Guglielmo, S., Monroe, A. E., & Malle, B. F. (2009). At the heart of morality lies folk psychology. *Inquiry*, 52(5), 449–466. <https://doi.org/10.1080/00201740903302600>
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852–870). Oxford University Press.
- Hardy, S. A. (2006). Identity, reasoning, and emotion: An empirical comparison of three Sources of moral motivation. *Motivation and Emotion*, 30(3), 205–213. <https://doi.org/10.1007/s11031-006-9034-9>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology Human Mental Workload*, 52, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10(3), 252–264. [https://doi.org/10.1207/s15327957pspr1003\\_4](https://doi.org/10.1207/s15327957pspr1003_4)
- Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Replication data for: Morality in everyday life. *Science*, 345(6202), 1340–1343. <https://doi.org/10.1126/science.1251560>
- Hsieh, S., & Lin, S. J. (2019). The dissociable effects of induced positive and negative moods on cognitive flexibility. *Scientific Reports*. <https://doi.org/10.1038/s41598-018-37683-4>
- Hsu, C.-F., Eastwood, J. D., & Toplak, M. E. (2017). Differences in perceived mental effort required and discomfort during a working memory task between individuals at-risk and not at-risk for ADHD. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2017.00407>
- Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. Appleton-Century.
- Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, 19(3), 126–132. <https://doi.org/10.1016/j.tics.2015.01.004>
- Inzlicht, M., Shenhav, A., & Olivola, C. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Science*, 22(4), 337–349. <https://doi.org/10.1016/j.tics.2018.01.007>
- Jordan, M. R., Amir, D., & Bloom, P. (2016a). Are empathy and concern psychologically distinct? *Emotion*, 16(8), 1107–1116. <https://doi.org/10.1037/emo0000228>
- Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016b). Third-party punishment as a costly signal of trustworthiness. *Nature*, 530(7591), 473–476. <https://doi.org/10.1038/nature16981>
- Jordan, J. J., & Kteily, N. (2020). Reputation drives morally questionable punishment. <https://doi.org/10.31234/osf.io/97nhj>
- Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2012). The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience*, 7(4), 393–402. <https://doi.org/10.1093/scan/nsr005>
- Kelley, H. H. (1971). *Attribution in social interaction*. General Learning Press.
- Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., Adolphs, R., & Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience*, 5(4), 378–384. <https://doi.org/10.1093/scan/nsq005>
- Kool, W., & Botvinick, M. (2013). The intrinsic cost of cognitive control. *Behavioral and Brain Sciences*, 36(06), 661–698. <https://doi.org/10.1017/s0140525x1300109x>
- Kool, W., & Botvinick, M. (2014). A labour/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, 143(1), 131–141. <https://doi.org/10.1037/a0031048>
- Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behavior*, 2(12), 899–908. <https://doi.org/10.1038/s41562-018-0401-9>
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, 139(4), 665–682. <https://doi.org/10.1037/a0020198>
- Kundu, P., & Cummins, D. D. (2013). Morality and conformity: The Asch paradigm applied to moral decisions. *Social Influence*, 8(4), 268–279. <https://doi.org/10.1080/15534510.2012.727767>
- Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, 7, 67–70. <https://doi.org/10.1016/j.copsyc.2015.08.003>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. NIMH, Center for the Study of Emotion and Attention.
- Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomatic major and corrugator supercilii. *Psychophysiology*, 40(5), 776–785. <https://doi.org/10.1111/1469-8986.00078>
- Leliveld, M. C., Dijk, E., & Van Beest, I. (2012). Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *European Journal of Social Psychology*, 42(2), 135–140. <https://doi.org/10.1002/ejsp.872>
- Ley, D. J. (2017, September 25). Coping with outrage fatigue: Are you exhausted with being upset? You're not alone. *Psychology Today*. <https://www.psychologytoday.com/ca/blog/women-who-stray/201709/coping-outrage-fatigue>
- Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 49(2), 433–442. <https://doi.org/10.3758/s13428-016-0727-z>
- MacCoun, R. J. (2005). Voice, control, and belonging: The double-edged sword of procedural fairness. *Annual Review of Law and Social Science*, 1, 171–201. <https://doi.org/10.1146/annurev.lawso.1.041604.115958>
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, 25(2), 147–186. <https://doi.org/10.1080/1047840X.2014.877340>
- Mikula, G., Scherer, K. R., & Athenstaedt, U. (1998). The role of injustice in the elicitation of differential emotional reactions. *Personality and Social Psychology Bulletin*, 24, 769–783. <https://doi.org/10.1177/0146167298247009>
- Miller, K. (2020, July 17). What is Doomscrolling? Experts explain why we do it—And how to stop. *Health*. <https://www.health.com/mind-body/what-is-doomscrolling>
- Mills, E. (2017, July 7). Take a break from the world and reduce your outrage fatigue syndrome. *USA Today*. <https://www.usatoday.com>

- [com/story/opinion/columnists/emily-mills/2017/07/07/mills-take-break-world-and-reduce-your-outrage-fatigue-syndrome/458879001/](https://www.nytimes.com/story/opinion/columnists/emily-mills/2017/07/07/mills-take-break-world-and-reduce-your-outrage-fatigue-syndrome/458879001/)
- Monroe, A. E., & Malle, B. F. (2017). Two paths to blame: Intentionality directs moral information processing along two distinct tracks. *Journal of Experimental Psychology: General*, *146*(1), 123–133. <https://doi.org/10.1037/xge0000234>
- Monroe, A. E., & Malle, B. F. (2019). People systematically update moral judgments of blame. *Journal of Personality and Social Psychology*, *116*(2), 215–236. <https://doi.org/10.1037/pspa0000137>
- Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind and Language*, *24*, 586–604. <https://doi.org/10.1111/j.1468-0017.2009.01375.x>
- Pizarro, D. A., & Tannenbaum, D. (2012). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 91–108). American Psychological Association.
- Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences of USA*. <https://doi.org/10.1073/pnas.2024292118>
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363. <https://doi.org/10.1037/1089-2680.7.4.331>
- Ronson, J. (2015, February 12). How one stupid tweet blew up Justine Sacco's life. *The New York Times Magazine*. <https://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html>
- Salerno, J. M., & Peter-Hagene, L. C. (2013). The interactive effect of anger and disgust on moral outrage and judgments. *Psychological Science*, *24*(10), 2069–2078. <https://doi.org/10.1177/0956797613486988>
- Scheffer, J. A., Cameron, D., & Inzlicht, M. (2021). Caring is costly: People avoid the cognitive work of compassion. <https://doi.org/10.1037/xge0001073>
- Silverman, H., & Ellis, R. (2019, June 30). A pregnant woman was shot, then indicted in her baby's death. Lawyers will try to get the charges dismissed. *CNN*. <https://www.cnn.com/2019/06/30/us/alabama-pregnant-woman-lawyers-motion/index.html>
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, *4*(4), 267–281. <https://doi.org/10.1111/j.1751-9004.2010.00254.x>
- Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*(6), 895–917. <https://doi.org/10.1037/0022-3514.88.6.895>
- Son, J., Bhandari, A., & Feldmanhall, O. (2019). Crowdsourcing punishment: Individuals reference group preferences to inform their own punitive decisions. *Scientific Reports*. <https://doi.org/10.31234/osf.io/ph3wn>
- Spring, V. L., Cameron, C. D., & Cikara, M. (2018). The upside of outrage. *Trends in Cognitive Sciences*, *22*(12), 1067–1069. <https://doi.org/10.1016/j.tics.2018.09.006>
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251. <https://doi.org/10.1037/0033-2909.87.2.245>
- Stets, J. E., & Carter, M. J. (2006). The moral identity: A principle level identity. In K. A. McClelland & T. J. Fararo (Eds.), *Purpose, meaning, and action: Control system theories in sociology* (pp. 293–316). Palgrave MacMillan.
- Tannenbaum, D., Uhlmann, E. L., & Diermeier, D. (2011). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology*, *47*(6), 1249–1254. <https://doi.org/10.1016/j.jesp.2011.05.010>
- Thomas, E. F., & McGarty, C. A. (2009). The role of efficacy and moral outrage norms in creating the potential for international development activism through group-based interaction. *British Journal of Social Psychology*, *48*(1), 115–134. <https://doi.org/10.1348/014466608X313774>
- Vásquez-Rosati, A., Montefusco-Siegmund, R., López, V., & Cosmelli, D. (2019). Emotional influences on cognitive flexibility depend on individual differences: A combined micro-phenomenological and psychophysiological study. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2019.01138>
- Viki, G. T., Fullerton, I., Raggett, H., Tait, F., & Wiltshire, S. (2012). The role of dehumanization in attitudes toward the social exclusion and rehabilitation of sex offenders. *Journal of Applied Social Psychology*, *42*(10), 2349–2367. <https://doi.org/10.1111/j.1559-1816.2012.00944.x>
- Winkielman, P., & Gogolushko, Y. (2018). Influence of suboptimally and optimally presented affective pictures and words on consumption-related behavior. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2017.02261>
- Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of USA*, *102*(20), 7398–7401. <https://doi.org/10.1073/pnas.0502399102>
- Yarkoni, T. (2019). The generalizability crisis. <https://doi.org/10.31234/osf.io/jqw35>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.