It's not what you do, but what everyone else does: On the role of descriptive norms and subjectivism in moral judgment

Andrew E. Monroe[a,*], Kyle D. Dillon[b], Steve Guglielmo[c], Roy F. Baumeister[d]

[a] *Appalachian State University, USA*
[b] *Harvard University, USA*
[c] *Macalester College, USA*
[d] *Florida State University, USA*

ABSTRACT

How do people evaluate moral actions, by referencing objective rules or by appealing to subjective, descriptive norms of behavior? Five studies examined whether and how people incorporate subjective, descriptive norms of behavior into their moral evaluations and mental state inferences of an agent's actions. We used experimental norm manipulations (Studies 1–2, 4), cultural differences in tipping norms (Study 3), and behavioral economic games (Study 5). Across studies, people increased the magnitude of their moral judgments when an agent exceeded a descriptive norm and decreased the magnitude when an agent fell below a norm (Studies 1–4). Moreover, this differentiation was partially explained via perceptions of agents' desires (Studies 1–2); it emerged only when the agent was aware of the norm (Study 4); and it generalized to explain decisions of trust for real monetary stakes (Study 5). Together, these findings indicate that moral actions are evaluated in relation to what most other people do rather than solely in relation to morally objective rules.

1. Introduction

In 2008, the United States and the world entered one of the worst recessions since the 1930s. One of the principle causes for the recession (at least in the U.S.) was a pervasive pattern of financial firms misrepresenting the quality of struggling investments and then "betting" that the investments would fail. In the wake of the financial collapse, many members of the public were morally outraged at such widespread deception. At the same time, many in the financial industry defended the behavior as having unfortunate consequences in this case, but being commonplace in the industry and therefore permissible.

This mismatch in moral perception mirrors a longstanding debate in moral psychology and philosophy. In many cases people hold a common intuition that standards of moral behavior are objective. That is, actions can be considered right or wrong regardless of culture or what others may believe or do. In line with this view, some moral beliefs—for example, prohibitions against killing—are so strongly ingrained in human morality that people adjudge them as if they were facts (Goodwin & Darley, 2012; Theriault, Waytz, Heiphetz, & Young, 2016). Contrastingly, some theorists argue that morality is largely subjective—that actions can only be morally evaluated according to the standards of the agent's culture. Considerable research on moral diversity supports this view, as many moral beliefs vary widely across, and even within, cultures (Graham, Haidt, & Nosek, 2009; Haidt & Graham, 2007; Turiel, Killen, & Helwig, 1987).

However, one limitation of most research on moral objectivism and subjectivism is that it often relies on cultural contrasts to derive evidence for either viewpoint (c.f., Goodwin & Darley, 2008). For example, Horvath and Giner-Sorolla (2007) demonstrate evidence for quasi-subjective moral judgment examining people's moral and legal judgments of extremely age-discrepant relationships (i.e., older men dating female minors). People exhibit strong moral and legal condemnation of these relationships, and specifically, of the senior partners in the relationship; however, this effect was moderated by whether both of the dating partners were from a country "where girls typically get married at 13". In this case, people sharply discounted blame and were reluctant to recommend severe legal punishment.

Some recent research suggests that moral-subjectivism can be studied without appealing to cross cultural differences. Recent developmental research suggests that people start out as moral objectivists (Nichols & Folds-Bennett, 2003), but they begin to endorse more subjectivist views as they age (Heiphetz & Young, 2017). Similarly Goodwin and Darley (2008) demonstrated that adults typically view morality as occupying a middle position between objective, factual

statements (e.g., the Earth is spherical) and subjective opinions.

More broadly, numerous accounts of moral judgment argue that human morality emerged to facilitate group functioning (Haidt, 2007; Malle, Guglielmo, & Monroe, 2012; Rai & Fiske, 2011), and therefore propose that many moral rules correspond to localized prescriptions regarding how members of a group should treat each other (Baumeister, 2005; Chudek & Henrich, 2011; Greene, 2013). Indeed behavior is often judged as worthy of blame or praise based on community taboos (Koleva, Graham, Iyer, Ditto, & Haidt, 2012), group values (Graham et al., 2011), and descriptive social norms (Sripada & Stich, 2006).

Thus, the present studies seek to examine whether people adopt moral-subjectivist patterns of judgment while holding culture constant. To test this possibility, we examine people's moral judgments of objectively negative and positive behaviors, but we vary the descriptive norms (i.e., what members of a community commonly do or believe) surrounding these behaviors. In this way, we examine whether descriptive norms are sufficient to explain moral-subjectivist patterns of judgment.

Several prominent perspectives imply that descriptive norms do not matter for people's moral judgments. Although the details of their accounts differ, many scholars suggest that moral judgments reflect considerations about actions themselves and about the consequences of actions (Conway & Gawronski, 2013; Cushman, 2013; Greene, 2013). For example, in the widely studied trolley dilemma, participants consider the moral permissibility of diverting a trolley from its current track, where it will kill five workers, to a side track, where it will kill one worker (Foot, 1967; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). People can morally disapprove of a behavior because they perceive the act itself to be inherently wrong (i.e., a deontological judgment) or because they perceive that the act has negative consequences (i.e., a consequentialist judgment). Notably, neither type of judgment should be sensitive to descriptive norms. Deontological judgments track whether an action was good/bad, while consequentialist judgments track how good/bad the action's consequences were; neither judgment tracks how the action relates to other people's typical behavior.

However, recent research suggests that even these types of sacrificial moral dilemmas are susceptible to descriptive norms and conformity effects. Bostyn and Roets (2017) demonstrate that people's endorsement of deontological or consequentialist moral decisions is strongly influenced by descriptive norms, whereby they favor the response they believe is consistent with the majority's view. Kundu and Cummins (2013) show similar results using an Asch conformity paradigm; participants reversed their decisions regarding which behaviors were (im)permissible to match the descriptively normative response in the room.

Descriptive norms provide information about what is common or expected within a group (Pettit & Knobe, 2009; Uttich & Lombrozo, 2010), and recognizing that a person has violated a norm of conduct is a first important step in activating people's moral judgments (Hitchcock & Knobe, 2009; Malle, Guglielmo, & Monroe, 2014). In fact, the emphasis on descriptive norms may reflect the process by which morality entered into human affairs. Tomasello (2016) proposed that morality began with dyadic concern and mutual obligation and only gradually superimposed cultural morals based on abstract principles on top of what he calls "second-person morality". Hence the basic mental structures for morality may have evolved initially to follow norms.

We therefore expect that descriptive norms will guide people's moral judgments, and we hypothesize two such pathways of influence. First, descriptive norms directly affect moral judgments by providing a behavioral standard that people use as a basis for judging a questionable act (Alicke, Rose, & Bloom, 2011; Fehr & Fischbacher, 2004). Second, descriptive norms affect moral judgments by revealing morally-relevant mental states (Guglielmo & Malle, 2010; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002; Uttich & Lombrozo, 2010). Moral acts often provide diagnostic information about an agent's desires and

intentions, and these mental state inferences, in turn, guide moral judgments about the agent (see Reeder, Monroe, & Pryor, 2008).

Five experiments tested the hypotheses that descriptive norms influence moral judgments both directly and indirectly (via mental state ascriptions). Study 1 sought to show that social norms regarding both positive and negative behavior shape moral judgments and mental state ascriptions. Study 2 disentangled the extremity of the agent's behavior from its norm-violating status by holding the agent's behavior constant and manipulating the content of the social norm. Study 3 provided a conceptual replication of Study 2's findings by comparing similar behaviors in different contexts—specifically tipping in the USA and UK, in which normative expectations differ. Study 4 assessed a potential boundary condition under which social norms may no longer influence moral judgment: when an agent lacks knowledge of the social norm. Last, Study 5 employed a behavioral economics paradigm with real monetary stakes to examine how descriptive norms shape a behavioral correlate of moral decision-making—namely, people's willingness to trust an interaction partner.

For all studies, we report all of our manipulations and measures. Each study's sample size was determined prior to data collection, and data analyses were always conducted following the completion of data collection. Materials and data for the experiments are available via OSF: https://osf.io/j5zyf/.

2. Study 1

Study 1 tested two predictions. First, that social norms would guide moral judgments for both positive and negative behavior: In both cases, exceeding a norm should elicit the most severe moral judgments (praise and blame), and falling below a social norm should elicit the weakest judgments. Second, that inferences about the agent's desires would mediate the relationship between norm-adherence and moral judgment. This prediction follows from research indicating that people use situational information to make inferences about the minds of others (Monroe & Reeder, 2011; Reeder et al., 2002, 2008), and that these mental state inferences, in turn, shape moral evaluations (Critcher, Inbar, & Pizarro, 2013; Inbar, Pizarro, & Cushman, 2012).

2.1. Method

2.1.1. Participants and procedure

We recruited 360 participants from Amazon Mechanical Turk (60 per condition). Sixteen participants were omitted from the analyses for failing to complete the study. Of the remaining participants, 61% were female. Average age in the sample was 32.8 years ($SD = 12.5$). All studies were approved by local IRBs. After data collection was complete, we conducted a sensitivity analysis with G*Power. The analysis showed that our sample size was sufficient to detect effect sizes of $\eta^2 = 0.028$ or larger with 80% power.

Participants were randomly assigned to one of six conditions in a 2 (valence: positive vs. negative) × 3 (norm condition: below, at, above) between-subjects design. They read a vignette that described an agent who performed either a morally positive behavior (donating to charity) or a morally negative behavior (cheating on his taxes), the magnitude of which differed relative to a known social norm.

2.1.2. Morally positive condition

John Smith lives in one of the 50 U.S. States. One morning, John sits down at his kitchen table to do his monthly paperwork. He notices a reminder about the upcoming charity drive.

John knows that most people donate about $4000 of their pay to charity. John donates [$1000/$4000/$7000] to charity.

2.1.3. Morally negative condition

John Smith lives in one of the 50 U.S. States. One morning, John sits down at his kitchen table to do his monthly paperwork. He notices a
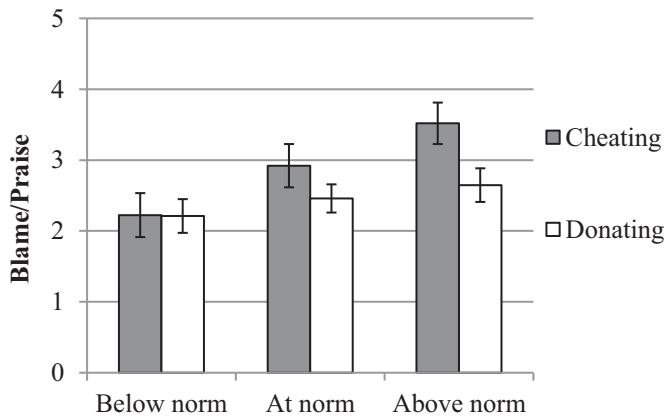
Fig. 1. Moral judgment extremity (praise and blame) across valence (cheating vs. donating) and norm conditions. Error bars = ± 1 SE.

reminder about paying his taxes.

John knows that most people cheat on their taxes by about $4000. John cheats on his taxes by [$1000/$4000/$7000].

Participants then rated how much blame or praise John deserved for his behavior (−5 [a lot of blame], 0 [neither praise nor blame], 5 [a lot of praise]) and how much John wanted to cheat/donate (1 [not at all] – 7 [very much]). In this and all subsequent studies we reverse scored participants' judgments of blame in the morally negative condition so that we could directly compare blame and praise ratings for the two behavior valence conditions. Lastly, participants completed a short demographic questionnaire.

2.2. Results

2.2.1. Norms, blame, and praise

A 2 (valence: positive vs. negative) × 3 (norm condition: below, at, above) ANOVA revealed a significant effect of norm condition, $F(2, 338) = 5.40$, $p = .005$, $\eta^2 = 0.031$, 95% CI [0.003, 0.07]. Exceeding the social norm amplified moral judgments of blame and praise relative to meeting the norm, and failing to meet the norm weakened judgments of blame and praise (See Fig. 1). There was also an effect of valence, $F(1, 338) = 4.27$, $p = .040$, $\eta^2 = 0.012$, 95% CI [0.00, 0.05], whereby judgments of blame were more extreme than judgments of praise. The effect of norm condition was not moderated by valence, $F(2, 338) = 1.32$, $p = .27$, $\eta^2 = 0.008$, 95% CI [0.00, 0.03].

2.2.2. Mental state ascriptions

Judgments of the agent's desires largely mirrored the pattern of moral judgments (See Fig. 2). Norm condition significantly influenced perceptions of desire, $F(2, 338) = 28.0$, $p < .001$, $\eta^2 = 0.14$, 95% CI
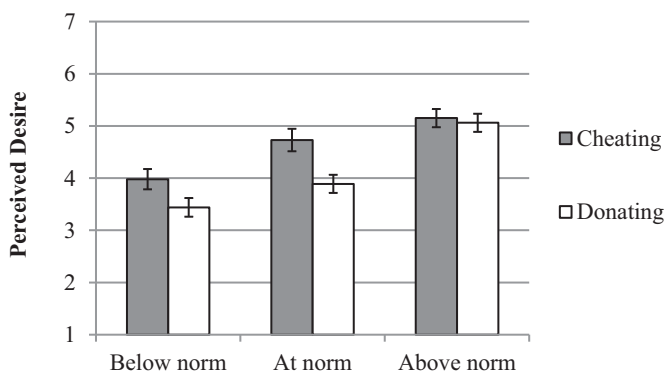

Fig. 2. Perceived desire across valence and norm conditions. Error bars = ± 1 SE.

[0.08, 0.21]. Perceived desire was strongest when the agent exceeded the norm, compared to when he matched or fell below the norm. There was also a main effect of valence, $F(1, 338) = 10.4$, $p = .001$, $\eta^2 = 0.03$, 95% CI [0.005, 0.07], such that participants attributed greater desire for cheating than for donating. Again, the norm by valence interaction was not significant, $F(2, 338) = 1.89$, $p = .15$, $\eta^2 = 0.01$, 95% CI [0.00, 0.04].

We then tested whether perceptions of desire mediated the relationship between norm condition and moral judgments using bootstrapping with 10,000 samples (Preacher & Hayes, 2008, model 4). As predicted, the effect of the norm condition on moral judgments (initial $b = 0.44$, $se = 14$, $p = .001$) was mediated (final $b = 0.12$, $se = 0.14$, $p = .37$) by perceived desire ($b = 0.44$, $se = 0.07$, $p < .001$), indirect $b = -0.31$, $se = 0.066$, 95% CI [0.195, 0.455]. Looking within each valence condition revealed an asymmetry in the strength of the mediation models. For negative behaviors, the effect of norms (initial $b = 0.65$, $se = 0.21$, $p = .003$) was mediated (final $b = 0.46$, $se = 0.21$, $p = .04$) by perceived desire ($b = 0.32$, $se = 0.12$, $p = .009$). By contrast, for positive behaviors, the norm condition did not significantly predict praise judgments, $b = 0.22$, $se = 0.16$, $p = .18$.

2.3. Discussion

Study 1 showed that experimentally manipulated descriptive norm information directly influenced people's moral judgments. As predicted, people graded their judgments of blame and praise based on how the target's behavior compared to a corresponding descriptive norm—whether the agent exceeded, matched, or fell below what other members of the social community were doing—rather than simply on whether the behavior violated an objective norm (cheating on taxes) or upheld a norm (donating to charity). Thus, this study shows that descriptive norms provide an important benchmark people use to grade the moral status of actions: People reduced blame and praise when an individual's behavior was less extreme than the social norm, and they ramped up blame and praise when the individual exceeded the norm.

Additionally, we found evidence for our prediction that norms shape moral judgments via inferences about agents' mental states. Overall, perceived desire mediated the relationship between norm condition and moral judgments. A closer analysis of this effect revealed an asymmetry for judgments of blame and praise. Perceived desires strongly mediated the relationship between norms and blame judgments but not praise judgments. Although this asymmetry was not predicted, it is consistent with previous work demonstrating that negative behaviors more easily capture attention, are more deeply processed, and more strongly predict moral judgments compared to positive behaviors (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Cushman, Dreber, Wang, & Costa, 2009; Pizarro, Uhlmann, & Salovey, 2003; Rozin & Royzman, 2001). Similarly, previous findings show that mental states more strongly predict negative moral judgments than positive ones because holding negative mental states is in itself blameworthy (Inbar et al., 2012; Woolfolk, Doris, & Darley, 2006) and because negative behaviors are perceived as clear indicators of mental states and dispositions, whereas positive behaviors may be multiply determined (Guglielmo & Malle, 2010; Reeder & Spores, 1983).

3. Study 2

Although Study 1 provided strong support for our hypotheses, one limitation is that it did not disentangle the effect of the norm information from the absolute magnitude of the behavior. For example, the above norm condition always presented the most extreme behavior (donating/cheating by $7000) and the below norm condition always presented the least extreme (donating/cheating by $1000). Thus, the results might simply reflect people's sensitivity to the absolute magnitude of the behavior (i.e., how much was cheated or donated) rather than their sensitivity to descriptive norms. To address this limitation

and disentangle the magnitude of the behavior from the descriptive norm, Study 2 held behavior constant while manipulating the descriptive norm.

Additionally, in this study we sought to replicate and explicitly test the positive-negative asymmetry for mental states mediating moral judgments. Study 1 demonstrated that although norm information informed mental state attributions for both negative and positive behaviors, mental states mediated the relationship between norms and moral judgments only for negative behaviors. This suggests that the positive-negative asymmetry may be explained in the linkage between the mental states and moral judgments. Past research clearly establishes that people view harboring harmful desires (even if they do not result in harm) as blameworthy (Cushman et al., 2009; Inbar et al., 2012; Woolfolk et al., 2006; Young & Saxe, 2009); however, the effect of praiseworthy desires on their own may not be as strong.

3.1. Method

3.1.1. Participants and procedure

260 participants (59% women, 78% white, 10% Black, 8% Latin/ Hispanic, average age 38.4, SD = 16.5) from the Tallahassee community were approached in a public park and asked to complete a brief psychological study. As in Study 1, we conducted a sensitivity analysis with G*Power. The analysis showed that, assuming 80% power, our sample size was sufficient to detect effect sizes of $\eta^2 = 0.036$ or larger.

After providing verbal consent, participants were randomly assigned to one of six conditions, in a 2 (valence: positive vs. negative) × 3 (norm condition: below, at, above) design. The valence conditions were identical to those in Study 1; however, here the agent's behavior was constant within each valence condition (i.e., cheating or donating $4000) and the social norm instead varied (cheating/donating by $1000/$4000/$7000).

3.1.2. Morally positive condition

John Smith lives in one of the 50 U.S. States. One morning, John sits down at his kitchen table to do his monthly paperwork. He notices a reminder about the upcoming charity drive.

John knows that most people donate about [$1000/$4000/$7000] of their pay to charity. John donates $4000 to charity.

3.1.3. Morally negative condition

John Smith lives in one of the 50 U.S. States. One morning, John sits down at his kitchen table to do his monthly paperwork. He notices a reminder about paying his taxes.

John knows that most people cheat on their taxes by about [$1000/ $4000/$7000]. John cheats on his taxes by $4000.

Following the story, participants rated how much blame or praise John deserved for his behavior (−5 [a lot of blame] – 0 [neither praise nor blame] – 5 [a lot of praise]) and how much John wanted to cheat/ donate (1 [not at all] – 7 [very much]). Participants then completed a short demographic questionnaire.

3.2. Results

3.2.1. Judgments of blame and praise

A 2 × 3 ANOVA revealed a significant effect of norm condition, $F(2, 254) = 3.93$, $p = .021$, $\eta^2 = 0.030$, 95% CI [0.004, 0.08], replicating the results from Study 1 (See Fig. 3). Judgments were most extreme when the agent exceeded the norm, and less so when the agent's behavior matched or fell below the norm. Additionally, we replicated the main effect of valence on moral judgments, $F(1, 254) = 10.78$, $p = .001$, $\eta^2 = 0.041$, 95% CI [0.006, 0.10], such that blame for negative behaviors was overall stronger than praise for positive behaviors. The norm by valence interaction again was not significant, $F(2, 254) = 0.496$, $p = .61$, $\eta^2 = 0.004$, 95% CI [0.00, 0.03].
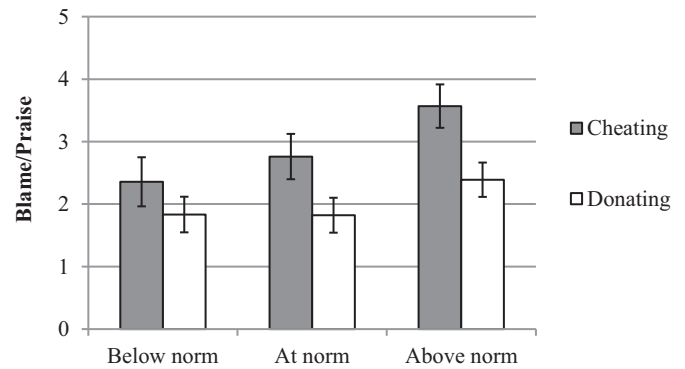
Fig. 3. Holding behavior constant, norms guide the extremity of praise and blame judgments. Error bars = ± 1 SE.

3.2.2. Perceived desires and moral judgment

Study 1 showed an asymmetric pattern of mediation, whereby perceived desire mediated the effect of norms on moral judgments for blame but not for praise. Accordingly, for Study 2 we used a moderated mediation model (Hayes, 2013, model 14), bootstrapping with 10,000 samples. Norm information significantly predicted perceived desires (See Table 1); however, when entered into the full model simultaneously with perceived desires, perceived desires remained a significant predictor of moral judgments ($p < .001$), while the direct effect of norms on moral judgments was only marginally significant ($p = .08$). Further, whereas perceived desire weakly predicted praise for positive behaviors, $b = 0.13$, $se = 0.06$, 95% CI [0.04, 0.27], it strongly predicted blame for negative behaviors, $b = 0.22$, $se = 0.09$, 95% CI [0.07, 0.55], resulting in significant moderated mediation (indirect $b = 0.09$, $se = 0.06$, 95% CI [0.006, 0.275], See Table 1 for model pathway coefficients).

3.3. Discussion

Study 2 conceptually replicated the findings of Study 1 while disentangling the effect of norm information from the absolute magnitude of the behavior. In Study 1, the norm was constant and the agent's behavior varied. Study 2 held the agent's behavior constant (donating/ cheating $4000) and instead varied the corresponding norm. The results again demonstrated that people's moral judgments were sensitive to the comparison between a behavior and its respective descriptive norm rather than merely to the magnitude of the outcome (i.e., the

Table 1
Unstandardized regression coefficients estimating perceived desire on moral judgments of blame and praise.

	Perceived desire (M)		Moral judgment (Y)	
	b (se)	95% CI	b (se)	95% CI[*]
Norm (X)	0.294** (0.111)	0.076, 0.513	0.267[+] (0.153)	−0.035, 0.570
Perceived desire (M)			0.589*** (0.086)	0.419, 0.758
Behavior valence (W)			0.342 (0.380)	−0.407, 1.091
X × W			−0.158[+] (0.085)	−0.325, 0.010
Constant	4.252*** (0.089) $R^2 = 0.026$ $F(1, 258) = 7.04, p = .008$	4.077, 4.429	−0.078 (0.385) $R^2 = 0.227$ $F(4, 255) = 18.74, p < .001$	−0.836, 0.680

[+] $p < .10$.
[*] $p < .05$.
[**] $p < .01$.
[***] $p < .001$.

absolute amount of cheating or donating). Thus, it is the relationship to the norm, and not the objective extremity of the behavior, that produced the different judgments.

Furthermore, Study 2 explicitly tested whether the effect of perceived mental states on moral judgments was moderated by behavior valence. Our findings replicated the blame-praise asymmetry from Study 1: Perceptions of a target's desires differentially mediated the relationship between descriptive norms and moral judgments as a function of the valence of the behavior (and thus the judgment) in question. Moreover, our tested model highlights how valence moderates the effect of mental state information on moral judgments. Specifically, our model demonstrates that people differentially utilize mental state information in their moral judgments of blame and praise. For both positive and negative behaviors, perceivers extracted information about an agent's desires, but those mental states subsequently influenced judgments of blame more strongly than praise.[1] This finding is consistent with research arguing that although perceived mental states do not always influence moral judgments (Young & Tsoi, 2013), they are especially important for judgments of blame (Griffin & Lombrozo, 2015).

4. Study 3

Studies 1 and 2, in presenting participants with explicit norm-based information regarding typical amounts of cheating and donating, revealed a consistent link between norms and moral judgments. Study 3 aimed to determine whether a similar pattern would emerge in the context of a more naturally occurring behavioral norm. Norms for tipping of service workers vary widely across cultures. For example, in the US, the local social norm is to tip 15% for dinner service, while in the UK it is 10%. By examining how these naturally varying norms affect people's moral appraisals of agents, we can more naturally disentangle the effect of norm-violating from the extremity of an agent's behavior. If descriptive norms play an important role in moral judgments, then falling below the local norm should be evaluated most negatively (regardless of the absolute size of the tip). This leads to the prediction that while participants in the UK might praise a person for tipping 10% (because it meets the UK's social norm), people in the US would blame the very same behavior (because it fails to meet the social norm in the US). The magnitude of the behavior is the same in both cases (giving someone a 10% tip) but the local norms differ. Our hypothesis was that identical behaviors would be judged differently based on their relation to descriptive norms.

4.1. Method

4.1.1. Participants

The US sample (*n* = 257) was recruited from Amazon Mechanical Turk. The UK sample (*n* = 242) consisted of volunteer participants at two UK universities: University of Kent and University College, London. A G*Power sensitivity analysis demonstrated that, assuming an 80% power level, our sample size was sufficient to detect effect sizes of $\eta^2 = 0.025$ or larger.

4.1.2. Procedure

Participants were randomly assigned to read one of four stories about a man who has dinner at a restaurant. The stories manipulated the size of the tip (5%, 10%, 15% or 20%) the man left his waiter. In the US, it is customary to tip 15% for dinner service, while in the UK it is

customary to tip 10%. Thus, in the US, tipping 10% is tipping below the norm; 15% is at the norm, and 20% is above the norm. Meanwhile, in the UK, tipping 5% is tipping below the norm; 10% is at the norm, and tipping 15% is above the norm. We focus on these six conditions.[2]

After reading the tipping story, participants rated how much blame or praise the character deserved (− 5 [a lot of blame], 0 [neither praise nor blame], 5 [a lot of praise]) and answered a question about subjective tipping norms ("What percentage of a bill do you think people should tip when dining out at a restaurant?").

4.2. Results

4.2.1. Endorsement of different tipping norms

We first verified that people in the US and UK endorsed different tipping norms. The data confirmed that UK participants believe that people should tip around 10% (*M* = 9.99, *SD* = 3.97), and US participants believe that people should tip around 15% (*M* = 16.34, *SD* = 4.55). These two tipping norms differed significantly, *t* (488) = 16.4, *p* < .001, *d* = 1.48, 95% CI [1.28, 1.68].

4.2.2. Tipping norms and moral judgments

A 3 (behavior: below, at, or above norm) × 2 (culture: US vs. UK) ANOVA tested the effect of norms on moral judgments. The levels of the behavior factor were implemented differently in the two samples: "below norm" was 5% in the UK but 10% in the US, "at norm" was 10% in the UK and 15% in the US, and "above norm" was 15% in the UK and 20% in the US. Results showed a significant effect of the behavior manipulation, *F*(2, 366) = 52.1, *p* < .001, $\eta^2 = 0.22$, 95% CI [0.15, 0.29] (see Fig. 4). Independent of the absolute magnitude of the tip, people's moral judgments were most negative when the agent failed to meet the norm and most positive when the agent exceeded the norm. This effect is vividly illustrated for tips of 10%. A post hoc test demonstrated that UK participants saw this behavior (i.e., meeting the norm) in relatively positive terms (*M* = 1.52, *SD* = 1.38), whereas US participants viewed the identical behavior of tipping 10% (i.e., falling below the norm) as morally blameworthy (*M* = −1.00, *SD* = 1.98), *t* (121) = 8.91, *p* < .001, *d* = 1.47, 95% CI [1.07, 1.87]. Overall, UK participants made more positive moral evaluations of the target than did US participants, *F*(1, 366) = 36.4, *p* < .001, $\eta^2 = 0.09$, 95% CI [0.04, 0.15], but there was no culture by norm interaction, *F*(2, 366) = 1.10, *p* = .33, $\eta^2 = 0.006$, 95% CI [0.00, 0.02].

4.3. Discussion

These data show that norm information plays an important role in people's moral evaluations of praise and blame. Moral evaluations were most positive when the person exceeded the local norm, and most negative when a person failed to at least meet the norm. These patterns emerged across both cultural groups. Further, the finding that UK participants made consistently more positive moral evaluations compared to US participants is striking given that, at every level of the norm manipulation, UK participants evaluated a lower absolute tipping amount compared to US participants (e.g., 5% vs. 10% tipping in the below norm condition).

Together, the findings from Studies 1–3 suggest that norms exert a consistent influence on people's moral evaluations. But in whose mind—the agent's, the perceiver's, or both—does the norm need to exist for it to be relevant for moral judgment? We examine this question in Study 4.

---

[1] We ran an additional post hoc mediation model, which allowed us to test for simultaneous moderation on the *a* and *b* pathways (Hayes, 2013, model 58). This model again showed the predicted moderated mediation of the b pathway (*p* = .06); however, there was no evidence of moderation of the *a* pathway (*p* = .74). In other words, norms equivalently influenced mental state inferences across valence, but mental states differentially influenced moral judgments.

[2] Including all eight conditions does not affect our pattern of results; however, we elected to constrain our analyses to testing our theoretically derived hypotheses. See Supplementary materials for analysis of all eight conditions.
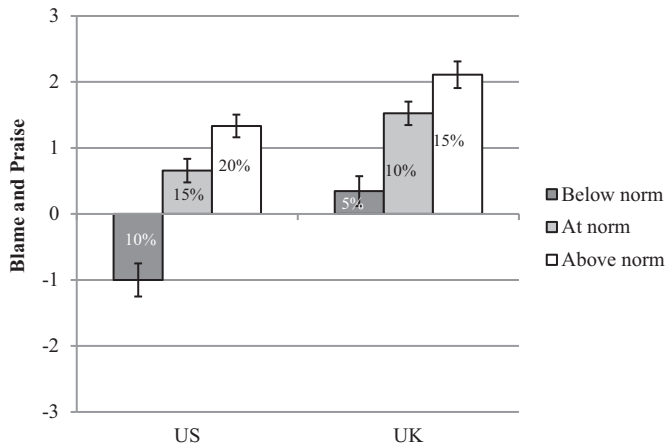
Fig. 4. Blame and praise ratings for tipping behavior by norm condition plotted across US and UK participants. Negative numbers indicate blame, whereas positive numbers indicate praise. Absolute tipping amounts are labeled on each bar. Error bars = ± 1 SE.

5. Study 4

In Studies 1–3, people reduced blame for an agent who acted less negatively than the norm. But is this because the agent knew his behavior fell below the norm, or because participants knew this? One possibility, supported by the notion of naïve realism (Ross & Ward, 1996), is that as long as perceivers (i.e., participants) are aware of the norm, it will influence their moral judgments of other agents' behaviors. This leads to the prediction that praise and blame will track norm information regardless of the agent's knowledge of the norm: For example, John's acting less negatively than the norm will elicit less blame, even when John doesn't know that his behavior falls below the norm. This pattern would be consistent with findings of outcome bias or moral luck, in which agents receive substantial blame or praise for the actual outcome of their behavior, regardless of their knowledge or intentions (Alicke, 1992). An alternative possibility, however, is that blame depends on the agent's own knowledge (e.g., Cushman, 2008; Young & Saxe, 2009). If an agent knows the social norm when he decides to act, then the agent's decision to exceed, match, or fail to meet the norm better illustrates his intentions in relation to the norm, which thereby gives perceivers reason to moderate blame and praise accordingly (Malle et al., 2014). In contrast, if an agent is ignorant of the norm when deciding how to act, then moral judgments should show little sensitivity to norm information because the agent acted without knowledge of these norms.

5.1. Method

5.1.1. Participants and procedure

A community sample ($N = 780$) from New England was recruited on local public beaches. Researchers approached potential participants and asked if they would be willing to complete a short survey. A G*Power sensitivity analysis demonstrated that, assuming an 80% power level, our sample size was sufficient to detect effect sizes of $\eta^2 = 0.012$ or larger.

After verifying that the participant was at least 18 years of age, the researcher left the participant to complete the study in private, returning to collect the data and to thank and debrief the participant. Participants were randomly assigned to one of twelve conditions, reflecting our previous valence (positive vs. negative) by norm condition (below vs. at vs. above) design, now crossed with a third factor: norm knowledge. In the norm known condition, the agent ostensibly knows the community norm of cheating or donating (e.g., "John Z. lives in a society where people donate about $4000 to charity"). By contrast, in

the norm unknown condition, the agent is aware of the presence of a social norm, but is ostensibly ignorant about its exact nature (e.g., "John Z. lives in a society where people donate some of their pay to charity"). Participants, however, learn the exact nature of the norm (e.g., "…it turns out that most people actually donate about $4000 of their pay to charity").[3]

After reading the story, participants rated how much blame or praise the agent deserved (−5 [a lot of blame], 0 [neither praise nor blame], 5 [a lot of praise]) and whether they believed the agent was aware of the norm when he acted ("Do you think that John knew that most people donate about $4000 to charity?" Yes/No).

5.2. Results

Confirming the effect of our knowledge manipulation, the majority of participants in the norm known condition (87%) reported that they believed that John was aware of the norm when he acted. By contrast, in the norm unknown condition fewer than half of participants (45%) believed that John was aware of the norm when he acted, χ^2 (1, N = 780) = 145.0, $p < .001$. In all subsequent analyses, we opted for a more conservative test of our hypothesis and report data from all participants; however, omitting participants based on their responses to the knowledge question does not alter our pattern of results.

Using a 2 (norm knowledge: known vs. unknown) × 3 (norm condition: below, at, above) × 2 (behavior valence: positive vs. negative) ANOVA we tested whether the effect of norm information on moral judgments of praise and blame would be moderated by the agent's knowledge of the norm. Results showed a significant knowledge by norm interaction, $F(2, 767) = 7.85$, $p < .001$, $\eta^2 = 0.02$, 95% CI [0.004, 0.042] supporting the hypothesis that an agent's knowledge of a norm is necessary for norm information to affect moral judgments. When the agent was aware of the norm, people graded their moral judgments based on whether his behavior fell below, met, or exceeded the norm, $F(2, 370) = 27.8$, $p < .001$, $\eta^2 = 0.13$, 95% CI [0.07, 0.19], consistent with the results of Studies 1 and 2. In contrast, when the agent was unaware of the norm, there was no effect of norm condition on moral judgments, $F(2, 403) = 0.18$, $p = .84$, $\eta^2 = 0.001$, 95% CI [0.00, 0.01] (see Fig. 5).[4]

5.3. Discussion

Study 4 indicated that norms influence moral judgments most clearly when the agent can be assumed to be aware of the norm. Thus, when assigning blame and praise, people considered the relationship between John's act and the corresponding behavioral norm, but only if John actually knew the norm. One important caveat to this finding, however, is that our manipulation of agent knowledge is quite subtle (in order to avoid demand characteristics), and a substantial set of participants reported that they believed John knew the norm in the norm unknown condition. Follow up analyses excluding these participants were consistent with our reported results (see Supplementary materials); however, we caution against overly strong interpretations of these data.

A possibly puzzling feature of Study 4's results was that condemnation was generally higher in the unknown norm condition than in the known norm condition (though not for the worst behaviors, that exceeded the norm of cheating). That is, people blamed John more for

---

[3] To minimize demand characteristics, neither condition included any explicit reference to John's knowledge or awareness of the norm.

[4] In addition to the predicted knowledge by norm interaction, the analysis revealed a main effect of norm condition ($F[2, 767] = 10.8$, $p < .001$, $\eta^2 = 0.027$, 95% CI [0.008, 0.053]) and valence, $F(1, 767) = 22.3$, $p < .001$, $\eta^2 = 0.028$, 95% CI [0.01, 0.055], as well as a significant knowledge by valence interaction, $F(2, 767) = 9.23$, $p = .002$, $\eta^2 = 0.023$, 95% CI [0.006, 0.04]. The three-way interaction was not significant, $F(2, 767) = 0.99$, $p = .37$, $\eta^2 = 0.003$, 95% CI [0.00, 0.01].
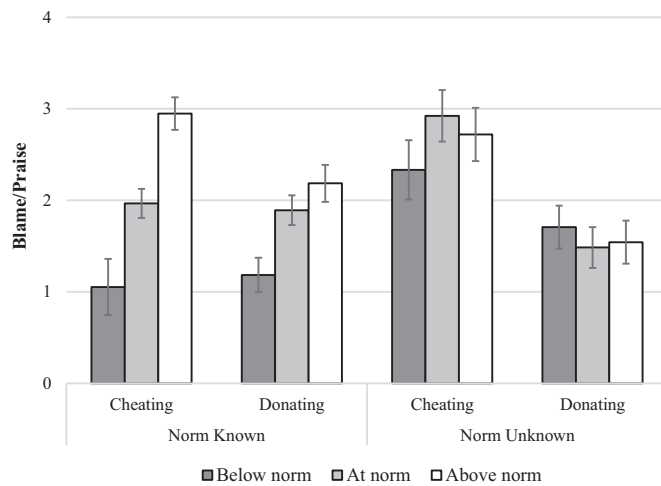
Fig. 5. Social norms of behavior influence moral judgments of praise and blame, but only when the agent is aware of the norm. Error bars = ± 1 SE.

cheating on his taxes when he was unaware of the norm than when he knew the norm. These findings therefore suggest that people are willing to grant leniency to agents who act less negatively than the prevailing social norm (i.e., cheating less than others), but only when agents know that they are acting in this way. After all, cheating is wrong, but if John knowingly cheats less than other people, he may be seen as acting based on somewhat socially acceptable grounds. People might interpret this act as deciding to cheat less than others and thereby reduce their condemnation for this decision. In contrast, if he does not know what others do, John's act is simply deciding to cheat, and he cannot justify his behavior by saying he chose to cheat less than others. In other words, his mental state in the norm unknown condition was that he clearly chose to do something bad, whereas in the norm known condition, he chose to do something less bad than most people. This fits the view that principled moral evaluation is an advanced form of judgment superimposed on the more basic and earlier form emphasizing the importance of group consensus and shared understandings.

6. Study 5

Studies 1–4 demonstrated that social norms and corresponding mental-state inferences play a consistent role influencing people's moral judgments. In particular, moral judgments were heightened when an agent exceeded a norm and weakened when an agent fell below a norm. Moreover, this differentiation was largely accounted for by perceptions of the agent's desire, and it emerged only when the agent was aware of the norm. Up to this point, our studies have focused on people's moral judgments in the context of hypothetical, third-person scenarios. Study 5 examined whether the relevance of norm information extends to contexts with more direct, tangible consequences for participants. In particular, this study used a paradigm from behavioral economics (the investment game) to probe the effect of social norm information on people's decisions about how much to trust an interaction partner.

6.1. Method

6.1.1. Participants and procedure

73 participants (42 female; mean age 19.7, $SD = 1.91$) from the Appalachian State University Psychology undergraduate subject pool were recruited for a study on "Economic Games" and run in groups of 9–12. A G*Power sensitivity analysis demonstrated that, assuming an 80% power level, our sample size was sufficient to detect effect sizes of $\eta^2 = 0.015$ or larger.

Upon entering the lab, participants were informed that they would

play a series of one-shot economic games with the other people in the room for a chance to win a cash bonus. The "get to know you" survey contained the demographic questionnaire for the study along with five questions about participants' recent behaviors and habits. The survey asked people to report "How many times over the last month have you told a lie to a friend (including minor or "white lies")?" and "How many times over the last month have you gone out of your way to recycle trash that wasn't yours?". To avoid raising suspicion about these two questions, they were embedded with three other questions asking about exercising habits, sleeping patterns, and current favorite TV interest.

Following participants' completion of the "get to know you" survey, the experimenter explained that participants would play six single-shot investment games with six different randomly selected people in the room. Participants were further told that some of them would be randomly selected to be Player 1 for all six games, while others would be randomly selected to be Player 2 for all six games. In reality, however, all participants were assigned to the role of Player 1. The experimenter explained that Player 1 would start each game with $10 and must decide how much (if anything) to send to Player 2. Whatever Player 1 sent would be tripled, and then Player 2 would decide how much money to send back to Player 1. Thus, the structure of the investment game is such that Player 1 maximizes their payoff by sending a larger amount of money to Player 2, but only to the extent that Player 2 can be trusted to send much of it back. Participants were informed that all of the games would be anonymous, and that they would not find out the results of any of the games until the end of the experiment. Participants were told that, at the end of the experiment, one of the games would be randomly selected, and participants would be paid a real cash bonus based on the results of that game.

Importantly, participants were told that although each of the games would be anonymous, they would receive three pieces of information about their partner that they could use to make their decisions: (1) their partner's response to either the "lying" or the "trash-pickup" question from the get to know you survey; (2) the group average for the corresponding question, and (3) their partner's ID number. We manipulated the information about the partner's "lying" or the "trash-pickup" question relative to the group average. Specifically, each participant played a game with a partner who exceeded, met, and fell-below both a negative norm (lying to friends) and a positive norm (picking up trash), resulting in six trials total (trials were presented in a randomized order for each participant). After participants completed the six investment game trials, they answered a question probing their suspicion of the study and its purpose and were subsequently paid and debriefed.

6.2. Results and discussion

A 2 (valence: lying vs. trash pickup) × 3 (norm: below, at, above) within-subjects ANOVA tested whether people utilize norm information in their moral decision-making—namely how much they trusted an interaction partner, operationalized as the amount of money they entrusted to Player 2. The analysis revealed a significant effect of valence, $F(1, 72) = 28.7$, $p < .001$, $\eta^2 = 0.28$, 95% CI [0.12, 0.43], whereby people entrusted less money when learning about a partner's negative behavior (lying) as compared to a partner's positive behavior (picking up trash). There was no overall effect of the norm manipulation, $F(2, 144) = 0.90$, $p = .41$, $\eta^2 = 0.01$, 95% CI [0.00, 0.05], but there was a powerful valence by norm interaction, $F(2, 144) = 80.7$, $p < .001$, $\eta^2 = 0.53$, 95% CI [0.41, 0.61] (See Fig. 6). In the positive norm condition (picking up trash), participants invested most in their partners when they exceeded the group norm, and less so when partners merely met or fell below the group norm ($ps < 0.001$, $\eta^2 = 0.28$ and 0.23 respectively). In contrast, in the negative norm condition (lying to friends) participants trusted least when their partner exceeded the group norm, and slightly more when partners met or fell below the group norm ($ps < 0.001$, $\eta^2 = 0.29$ and 0.35 respectively).

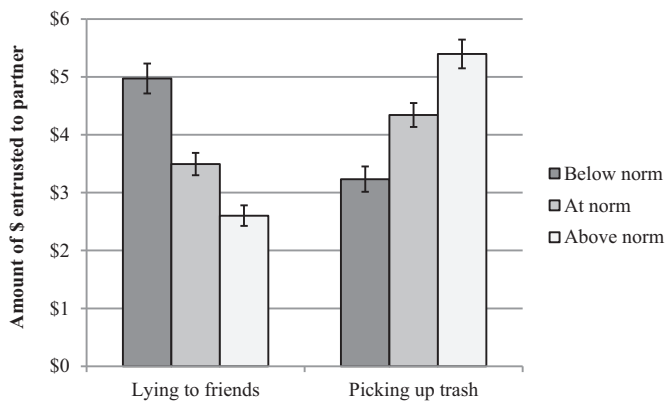Importantly, these effects held after omitting participants who

Fig. 6. Participants' decisions to entrust a partner with money were strongly influenced by a partner's norm-relevant behavior. Error bars = ± 1 SE.

expressed even minimal suspicion about the study (e.g., wondering if all of the partner matchups were real). While this exclusion reduced the sample size substantially (resultant $n = 39$), the significant valence x norm interaction remained strong, $F(2, 76) = 40.5$, $p < .001$, $\eta^2 = 0.52$, 95% CI [0.35, 0.62], as did the effect of valence, $F(1, 38) = 19.2$, $p < .001$, $\eta^2 = 0.34$, 95% CI [0.11, 0.52]. The main effect of norm information also remained non-significant $F(2, 76) = 1.02$, $p = .36$, $\eta^2 = 0.03$, 95% CI [0.00, 0.11].

Thus, the moral judgment findings of Studies 1–4 are not merely abstract opinions about hypothetical characters. Study 5 required participants to make consequential decisions involving real money. Their inclination to trust someone depended not only on whether the person's behavior was good or bad but on how it compared to the group's descriptive norms.

7. General discussion

Five studies provided evidence that when people render moral judgments of blame and praise, they do so by referencing the descriptive norms of a community rather than by solely invoking an objective norm. In order to decide how much blame or praise a person deserves, perceivers evaluate how that person's behavior compares to what is common in the community. Although people often think of moral judgments as objective, our findings indicate another factor that influences judgments of morality: what others in the community do or accept as commonplace. Moral norms may thus have a descriptive as well as an objective component. The influence of descriptive norms on moral judgments held true for moral judgments of blame and praise (Studies 1–4) as well as for decisions about trusting someone with money (Study 5). Moreover, this effect was consistent even when holding the agent's behavior constant and varying the descriptive norm (Studies 2–3).

Importantly, however, our data suggest that agents must be aware of a descriptive norm for it to impact perceivers' moral judgments (Study 4). Behaviors were not judged differently based on their relationship to descriptive norms of which the actors were unaware. These data indicate that moral judgments focus on the mental state of the person being judged, and not just the action itself. A revealing twist in the data added further evidence of this. In Study 4, being ignorant of the descriptive cheating norm did not reduce blame. On the contrary, tax cheaters were condemned even more when they did not know the descriptive norm compared to when they knew them (though this effect appears only for mild and moderate cheating). More concretely, an agent who cheats on his taxes by $1000 gets some credit (diminished blame) if he knows that most other people cheat by $4000, because he is (presumably) choosing to be more virtuous than others. But if he cheats by $1000 while not knowing that the average person cheats by $4000, he gets no credit for virtuous restraint. He is simply

intentionally cheating on his taxes.

Norms influenced not only perceivers' judgments of blame, praise, and desire but also their behavior toward others in a context with real monetary consequences (Study 5). Thus, our findings apply beyond judgments of hypothetical actors to additionally illuminate the way people use descriptive norm information to make real and potentially costly decisions. It may seem surprising that people trusted more of their money to someone who admitted lying to friends than someone who reported helping the environment by recycling—but, this finding highlights the power of subjectivist reasoning in the context of descriptive norms of behavior. Not knowing anything else, people would rather trust someone who is less dishonest than their neighbors compared to someone who is virtuous, but less virtuous than the other people in the room.

The current findings might also provide an interesting lens through which to understand recent political events. In 2016, then-candidate Donald Trump was widely acknowledged as a person who made misleading or factually incorrect statements; however, at the same time, a common view in the public was that "all politicians were liars". Trump's tendency to make off-color statements lent itself to a narrative that he might lie less than other politicians. Polling data the week before the 2016 election support this view. Potential voters viewed Trump as more trustworthiness than Hilary Clinton (46% vs. 38%, Washington Post-ABC, 2016) who had been dogged by accusations of being an especially dishonest politician. Our data explain part of these trends by arguing that even though it was commonly acknowledged that Trump was a liar, he was viewed as less of a liar than the norm in politics (and his political opponent). Thus, this perceived deviation from the norm may have been one factor (among others) that helped buoy him to the Presidency.

7.1. Blame-praise asymmetries

Our studies also revealed two blame-praise asymmetries. First, people, on average, blamed negative norm violations more extremely than they praised positive norm violations (Baumeister et al., 2001; Gneezy & Epley, 2014; Keysar, Converse, Wang, & Epley, 2008; Pizarro et al., 2003). These data are consistent with the view that negative events more easily attract people's attention and motivate processing (Baumeister et al., 2001; Rozin & Royzman, 2001), that people apply different standards to blame versus praise judgments (see Kessler et al., 2010), and that people tend to make more extreme blame judgments than praise judgments, even when outcomes are equated (Cushman et al., 2009).

Study 5 highlighted a potentially interesting deviation from this general pattern of findings. When deciding how much to trust an interaction partner, negative and positive norm information exerted similar effects on trusting behavior. One possible explanation, consistent with past research, is that people apply more stringent constraints on expressing negative moral judgments when they believe they will be publicly observable (see Ferber & Monroe, 2016; Monroe & Malle, under review; Tetlock, 1983). Whereas, Studies 1–4 demonstrate a blame-praise asymmetry for relatively private and perhaps low stakes moral judgments; when judgments are ostensibly public and stakes are high people's treatment of morally negative and positive norm information may become more even handed.

The second blame-praise asymmetry pertains to the role of mental states in informing moral judgments. Observers inferred greater desire from negative than from positive behaviors, but perhaps more importantly, this inferred desire more strongly mediated the relationship between norms and moral judgment for blame than for praise. Although we did not measure all steps in attributional reasoning, the pattern is consistent with general attributional logic. Some people perform virtuous acts not from personal motivation but from a sense of obligation or social pressure. In contrast, there is no social pressure or obligation to cheat — if anything, such immoral actions go against social

pressures, which intensifies the impression that the person really wanted to act that way. Thus, in the mind of perceivers, an agent's positive behavior may be multiply determined, thereby providing ambiguous information about the agent's true motivation, whereas negative behavior is usually viewed as a clear indicator of mental states and dispositions (Reeder & Spores, 1983).

7.2. Concluding remarks

Moral beliefs are often regarded as objective, both by experts and laypersons. Our findings do not discredit that view, but they modify it. The present studies show that people judge behavior not only by objective standards, but also by how it conforms to or deviates from subjective, descriptive norms. In this way, our findings suggest that people view morality as situated between pure objectivism and subjectivism, consistent with previous research (Goodwin & Darley, 2012). Immoral behavior is ubiquitously blamed, but the specific degree of blame depends on whether a behavior exceeds or falls short of the descriptive norm. Similarly, praise for virtuous behavior is pervasive, but the amount of praise people award to agents is shaped—albeit more weakly than for blame—by subjective, descriptive norms.

The key role of descriptive norms dovetails with the view that morality is, in substantial part, a system for facilitating harmonious and cooperative communities (e.g., Carnes, Lickel, & Janoff-Bulman, 2015; Haidt, 2007; Rai & Fiske, 2011). Norms were incorporated into group life as a way of codifying what the group regarded as proper. But since people will almost never adhere perfectly to all such norms, what matters for securing one's inclusion in a group is to be good enough—and this depends not just on adherence to a standard but on calibrating one's behavior to what is commonplace in one's group.

Open practices

The experiment in this article earned Open Materials and Open Data badges for transparent practices. Materials and data for the experiments are available at https://osf.io/j5zyf/.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2018.03.010.

References

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology, 63*, 368–378. http://dx.doi.org/10.1037/0022-3514.63.3.368.

Alicke, M. D., Rose, D., & Bloom, D. (2011). Causation, norm violation and culpable control. *Journal of Philosophy, 108*, 670–696.

Baumeister, R. F. (2005). *The cultural animal: Human nature, meaning, and social life.* New York, NY, US: Oxford University Press.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370.

Bostyn, D., & Roets, A. (2017). An asymmetric moral conformity effect: Subjects conform to deontological but not consequentialist majorities. *Social Psychological and Personality Science, 8*, 323–330. http://dx.doi.org/10.1177/1948550616671999.

Carnes, N. C., Lickel, B., & Janoff-Bulman, R. (2015). Shared perceptions: Morality is embedded in social contexts. *Personality and Social Psychology Bulletin, 41*, 351–362. http://dx.doi.org/10.1177/0146167214566187.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15*, 218–226. http://dx.doi.org/10.1016/j.tics.2011.03.003.

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology, 104*, 216–235. http://dx.doi.org/10.1037/a0031021.

Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science, 4*, 308–315. http://dx.doi.org/10.1177/1948550612457688.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*, 353–380. http://dx.doi.org/10.1016/j.cognition.2008.03.006.

Cushman, F. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and Social Psychology Review, 17*, 273–292. http://dx.doi.org/10.1177/

1088868313495594.

Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a "trembling hand" game. *Plos One, 4*, e6699.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*, 63–87.

Ferber, L., & Monroe, A. E. (2016). *Seething but quiet: Power differentially affects public vs. private expressions of blame. Presented at the Society for Judgment and Decision Making, Boston, MA.*

Foot, P. (1967). *The problem of abortion and the doctrine of double effect. 5*, Oxford Review5–15.

Gneezy, A., & Epley, N. (2014). Worth keeping but not exceeding asymmetric consequences of breaking versus exceeding promises. *Social Psychological and Personality Science, 5*, 796–804. http://dx.doi.org/10.1177/1948550614533134.

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition, 106*, 1339–1366. http://dx.doi.org/10.1016/j.cognition.2007.06.007.

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology, 48*, 250–256. http://dx.doi.org/10.1016/j.jesp.2011.08.006.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029–1046. http://dx.doi.org/10.1037/a0015141.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366–385. http://dx.doi.org/10.1037/a0021847.

Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them.* New York: Penguin Press.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105–2108. http://dx.doi.org/10.1126/science.1062872.

Griffin, C., & Lombrozo, T. (2015). *Mental states are more important in evaluating moral than conventional violations.*

Guglielmo, S., & Malle, B. F. (2010). Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. *Personality and Social Psychology Bulletin, 36*, 1635–1647. http://dx.doi.org/10.1177/0146167210386733.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998–1002. http://dx.doi.org/10.1126/science.1137651.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98–116. http://dx.doi.org/10.1007/s11211-007-0034-z.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press.

Heiphetz, L., & Young, L. L. (2017). Can only one person be right? The development of objectivism and social preferences regarding widely shared and controversial moral beliefs. *Cognition, 167*, 78–90. http://dx.doi.org/10.1016/j.cognition.2016.05.014.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *The Journal of Philosophy, 106*, 587–612.

Horvath, M. A. H., & Giner-Sorolla, R. (2007). Below the age of consent: Influences on moral and legal judgments of adult–adolescent sexual relationships. *Journal of Applied Social Psychology, 37*, 2980–3009. http://dx.doi.org/10.1111/j.1559-1816.2007.00290.x.

Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting from misfortune when harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin, 38*, 52–62. http://dx.doi.org/10.1177/0146167211430232.

Kessler, T., Neumann, J., Mummendey, A., Berthold, A., Schubert, T., & Waldzus, S. (2010). How do we assign punishment? The impact of minimal and maximal standards on the evaluation of deviants. *Personality and Social Psychology Bulletin, 36*, 1213–1224. http://dx.doi.org/10.1177/0146167210380603.

Keysar, B., Converse, B. A., Wang, J., & Epley, N. (2008). Reciprocity is not give and take asymmetric reciprocity to positive and negative acts. *Psychological Science, 19*, 1280–1286. http://dx.doi.org/10.1111/j.1467-9280.2008.02223.x.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially purity) help explain culture war attitudes. *Journal of Research in Personality, 46*, 184–194. http://dx.doi.org/10.1016/j.jrp.2012.01.006.

Kundu, P., & Cummins, D. D. (2013). Morality and conformity: The Asch paradigm applied to moral decisions. *Social Influence, 8*, 268–279. http://dx.doi.org/10.1080/15534510.2012.727767.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2012). Moral, cognitive, and social: The nature of blame. In J. P. Forgas, K. Fiedler, & C. Sedikides (Eds.). *Social thinking and interpersonal behavior, Sydney symposium of social psychology* (pp. 313–331). New York, NY US: Psychology Press.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry, 25*, 147–186. http://dx.doi.org/10.1080/1047840X.2014.877340.

Monroe, A. E., & Malle, B. F. (2018). *Testing a socially-regulated blame hypothesis: People systematically update moral judgments of blame.* (under review).

Monroe, A. E., & Reeder, G. D. (2011). Motive-matching: Perceptions of intentionality for coerced action. *Journal of Experimental Social Psychology, 47*, 1255–1261. http://dx.doi.org/10.1016/j.jesp.2011.05.012.

Nichols, S., & Folds-Bennett, T. (2003). Are children moral objectivists? Children's judgments about moral and response-dependent properties. *Cognition, 90*, B23–B32. http://dx.doi.org/10.1016/S0010-0277(03)00160-4.

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language, 24*, 586–604. http://dx.doi.org/10.1111/j.1468-0017.2009.01375.x.

Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science, 14*, 267–272.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891. http://dx.doi.org/10.3758/BRM.40.3.879.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*, 57–75. http://dx.doi.org/10.1037/a0021867.

Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology, 83*, 789–803.

Reeder, G. D., Monroe, A. E., & Pryor, J. B. (2008). Impressions of Milgram's obedient teachers: Situational cues inform inferences about motives and traits. *Journal of Personality and Social Psychology, 95*, 1–17. http://dx.doi.org/10.1037/0022-3514.95.1.1.

Reeder, G. D., & Spores, J. M. (1983). The attribution of morality. *Journal of Personality and Social Psychology, 44*, 736–745.

Ross, L., & Ward, A. (1996). Naive realism: Implications for social conflict and misunderstanding. In E. S. Reed, E. Turiel, & T. Brown (Eds.). *Values and knowledge* (pp. 103–135). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*, 296–320. http://dx.doi.org/10.1207/S15327957PSPR0504_2.

Sripada, C. S., & Stich, S. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Eds.). *The innate mind (Vol. 2: Culture and cognition)* (pp. 280–301). New York, NY: Oxford University Press.

Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly, 46*, 285–292. http://dx.doi.org/10.2307/3033716.

Theriault, J., Waytz, A., Heiphetz, L., & Young, L. (2016). *The mind behind metaethics: Perceptions of moral objectivity and subjectivity rely on theory of mind*.

Tomasello, M. (2016). *A natural history of human morality.* Cambridge, MA: Harvard University Press.

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan, & S. Lamb (Eds.). *The emergence of morality in young children* (pp. 155–243). Chicago, IL, US: University of Chicago Press.

Uttich, K., & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition, 116*, 87–100. http://dx.doi.org/10.1016/j.cognition.2010.04.003.

Washington Post-ABC (2016). *Post-ABC tracking poll, Oct. 28–31, 2016*.

Woolfolk, R. L., Doris, J. M., & Darley, J. M. (2006). Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility. *Cognition, 100*, 283–301. http://dx.doi.org/10.1016/j.cognition.2005.05.002.

Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia, 47*, 2065–2072. http://dx.doi.org/10.1016/j.neuropsychologia.2009.03.020.

Young, L., & Tsoi, L. (2013). When mental states matter, when they don't, and what that means for morality. *Social and Personality Psychology Compass, 7*, 585–604. http://dx.doi.org/10.1111/spc3.12044.