Negative desires make failure to help more blameworthy:

The role of wrongness and moral character evaluations

Paula Yumi Hirozawa (Graduate School of Informatics, Nagoya University, yumi.hirozawa@gmail.com)
Minoru Karasawa (Graduate School of Informatics, Nagoya University, mkarasawa@nagoya-u.jp)

望ましくない願望の存在が援助不履行への非難を招く―道徳的人物評価が不作為に関する判断に与える影響―

ヒロザワ パウラ ユミ（名古屋大学 大学院情報学研究科）

唐沢 穣（名古屋大学 大学院情報学研究科）

要約

加害行為に対する非難は一般に、当該行為と結果の因果関係、およびその行為に関わる意図の程度に左右される。こうした因果性と意図性に基づく道徳的推論は、一般人の判断だけでなく法の専門家による判断においても同様の役割を果たす。しかし近年の研究は、行為により示唆される望ましくない願望の存在が、むしろ非難の程度に影響することを示している。本研究では、援助をし損ねた行為者の願望に関する情報が、非難判断に与える影響を調べた。実験シナリオには、他者を援助すると約束した行為者が、「望ましい」（約束通り援助したい）、「中立的」（援助したいともしたくないとも思わない）、「望ましくない」（約束した相手に悪意を抱き、援助したくない）のいずれかの願望を持っていたと記されていた。続いて、実際には援助がなされず約束相手に害が及んだことが示された後、行為者に対する非難の程度について評定が求められた。結果は、行為者の不履行が結果の直接的原因でなかったにもかかわらず、「望ましくない」願望の条件において特に強い非難が行われることを示した。さらにこの効果を、援助不履行に関する道徳判断と、行為者の道徳的人格評価が媒介していた。一方、行為と結果の因果性、および意図性に関する知覚の媒介効果は見られなかった。道徳判断における属人的判断、そして不作為に関する法的判断などの、より実際的な文脈に関する示唆について議論する。

Key words

desires, blame, wrongness, moral character, failure to help

1. Introduction

Imagine Nakamura promises he would take his friend to the airport. However, Nakamura ends up unable to keep his promise due to an urgent last-minute meeting at work and, as a consequence, his friend misses his flight. To what extent should Nakamura be blamed for the consequence that his friend missed his flight? Now, imagine that, despite the fact that his failure to help was not intentional, Nakamura admits he had no desire to help his friend and secretly wished his friend would have a hard time. Would the information of Nakamura's desires affect judgments of blame for his failure to help?

In person's folk conception, intentionality is structured of five components - desires, belief, intentions, awareness, and skill. Desires reflect what individuals wish and want, and desires by themselves do not imply action. For instance, a person may want to save the world but that is different from deciding to act on it (Malle & Knobe, 1997). Intuitively, desires are implied in the interpretation of intentional action. However, individuals also recognize that acts can be intentional without the component of desire, which implies that people distinguish between desires and intentionality (Nadelhoffer, 2006).

In both criminal experts' and layperson's moral reasoning, judgments of blame depend heavily on whether the agent had a causal role in the perpetrated harm and whether they commit-ted the harm intentionally (Cushman, 2008; Gray et al., 2012; Greene et al., 2009; Lagnado & Channon, 2008; Malle et al., 2014; Shaver, 1985; Woolfolk et al., 2006). However, recent research shows evidence that not always causality and intentionality are necessary for assigning blame. In Inbar et al.'s (2012) set of four studies, they found that individuals who benefited from a misfortune (e.g., winning a bet that a natural disaster will occur) were deemed more blameworthy for their acts compared with those who did not benefit from it, despite having no causal or intentional role in the disaster occurrence. The results were explained by underlying perceptions of negative, or "wicked" desires. Based on the person-centered approach to moral judgments (Pizarro &Tannenbaum, 2011; Uhlmann et al., 2015), the authors speculate that people are motivated to blame immoral agents, and hence benefitting from misfortune received greater blame because of an underlying assumption that only bad agents would have such wicked desires.

In this study, we examined whether the information of one's positive versus negative desires affect judgments of blame in the context of failure to help. Instances of blame for failing to help others are largely prevalent in ordinary life and extends to the legal sphere (e.g., judgments of omission), yet little research has so far focused on the moral reasoning behind such judgments. The well-evidenced omission effect refers to the tendency of judging commission (i.e., committed acts) more harshly than omission (i.e., failing to act in order to prevent the negative outcome) (Bostyn & Roets, 2016; Cushman, 2008; Ritov & Baron,

1990; Spranca et al., 1991). This effect occurs in part because omissions allow multiple inferences of causality (e.g., the cause of an accident can be assigned to the agent who failed to prevent it, but also to any other individual who also failed to prevent the outcome), hence attenuating blame (Bostyn & Roets, 2016; Spranca et al., 1991). Inferences of intentionality are also mitigated for omission (Hayashi, 2015). Whereas omissions are less blameworthy than commissions, it is nonetheless a point that they are still deemed blameworthy. Based on Inbar et al.'s (2012) findings, we believe the perceptions of an agent's desires is an important factor to understand the blaming process in such cases which do not involve direct causality and intentionality.

In this study, we described agents who promised to help another person yet unintentionally failed to help. We manipulated the extent to which the agents desired to help. In the "negative desire" condition, the agent does not desire to help and has negative wishes towards the promised individual. On the other hand, in the "positive desire" condition, the agent desires to help and has positive wishes towards the promised individual. Finally, there was a "neutral desire" condition, in which the agent does not particularly want to help but does not oppose it either. Considering previous evidence on the effect of desires on blame, we expected the agent with negative (vs. positive) desires to receive the greatest level of blame. Establishing a neutral condition also allowed us to examine whether information of positive desires mitigates blame.

In Inbar et al.'s (2012) study, blameworthiness was measured by a combined composite of blame and wrongness. However, these two judgments display distinct qualities. Judgments of wrongness are more sensitive to mental states information, such as beliefs and desires, whereas blame is determined by both mental state and outcome information (Cushman, 2008). Helping others and fulfilling one's social obligations are commendatory and normative behaviors (Janoff-Bulman et al., 2009) that are socially reinforced (Graham et al., 2013; Shweder et al., 1997). It seems people expect others not only to behave positively but to also present corresponding positive mental states. For instance, people seem to infer that individuals in general have good desires (Pizarro et al., 2003), and good intentions are not as diagnostic of character compared to negative intentions (Hirozawa et al., 2020), suggesting that positive mental states are deemed as default. In general, these findings suggest that people expect others to act with positive desires, which implicates that acting with negative desires may be perceived to be a wrongful course of action and, hence, more blameworthy. To examine this possibility, we treated wrongness as a potential mediator of the effect of desires on blame.

As previously noted, a second mechanism through which people may assign greater blame for failure to help with negative desires is based on person evaluations. The person-centered approach to moral judgments posits that blame is assigned with the motivation to blame immoral characters inferred by their mental state (Pizarro & Tannenbaum, 2011). Previous research shows evidence of a dissociation between moral judgments of acts and that of persons. For instance, bigots received greater blame compared to physical assailants, even though racial slur was perceived to be a less immoral violation than physical assault (Uhlmann et al., 2013). People were deemed blameworthy even when performing harmless acts, as long as their behavior seemed informative of a bad character (see Pizarro et al., 2012). Hence, people may determine blame for failure to help not only on the basis of the wrongness of the act but also on what the act informs about the person. We also examined the potential role of perceived moral character as a second mediator.

Finally, we elaborated the scenarios in a way that the perceptions of causality and intentionality should remain constant. That is, regardless of the desire condition, all agents failed to keep their promise unintentionally and for the same reasons. However, a motivated reasoning account suggests that mental state information can alter following factual moral judgments. For instance, an agent was perceived to be more causal of an accident when he was driving home to hide cocaine (vs. hide a gift for his parents) (Alicke, 1992). People also assigned greater harm to intentional (vs. unintentional) acts even when the harm was identical (Ames & Fiske, 2013). This account predicts that people may inflate moral judgments to justify their blame motivation. To address to this potential explanation, we also included measures of causality and intentionality.

In sum, in the present study, we examined whether individuals who unintentionally fail to help with negative desires (vs. positive and neutral desires) receive greater blame. We hypothesized judgments of wrongness of their failure to help and inferences of immoral character to be potential mediators of this effect. To account for a potential blame validation phenomenon, we also included causality and intentionality in the mediational analyses.

2. Method
2.1 Participants
We recruited 210 participants through an outsourcing service, CrowdWorks (112 females, M_{age} = 38.78, SD = 8.77). The sample size was determined by prior analysis using GPower. Due to the nature of the Latin Square design, we chose for a conservative calculation based on each pair of scenarios, that is, a between-subjects Analysis of Variance for main effects and interaction (η_p^2 = 0.09, α = .05, $1 - \beta$ = .80, df = 2, $groups$ = 6). The minimum sample size for a set of two scenarios was 101. We determined the sample size prior to any data analysis, and there was no exclusion of participants.

2.2 Materials, design, and procedures
We prepared six scenarios describing agents who failed to keep their promise to help another person, followed by a negative outcome. Specifically, the scenarios depicted agents who

failed to keep the promise to: (1) donate to their friend's project, which ended up getting cancelled for lack of donations; (2) buy a birthday cake for their friend, spoiling the birthday party; (3) help their cousin with moving out, resulting in the cousin having to pay a fee to the real state agency; (4) help a classmate with math, resulting in the classmate's failure in the test; (5) take care of their sick grandfather, who became more ill; (6) take a friend to the airport, who ended up missing their flight. In all scenarios, the agents made the promise compelled by either the promised person or a third person (e.g., the friend asked for a lift to the airport; the agent's mother asked the agent to help the cousin). The agents failed to help due to forgetfulness or an external justification (e.g., they realized they had no money when they were about to donate). We separated these scenarios into one set (A) containing scenarios 1, 2, and 3, and another set (B) containing scenarios 4, 5, and 6. Participants were randomly assigned to either one of these sets (see Appendix for full scenarios).

Moreover, as detailed in the introduction, we manipulated three levels of desires. For example, in the Airport scenario, the agent's desires were described as either:

- Positive:
  Deep inside, Hirata really wanted to drive the friend to the airport. S(h)e[1] thought friends should help each other and wished the friend would have a pleasant journey.
- Neutral:
  Deep inside, Hirata didn't particularly want to drive the friend to the airport but didn't mind doing it either. S(h)e felt neutral about the friend's request.
- Negative:
  Deep inside, Hirata really did not want to drive the friend to the airport. S(h)e thought the friend was inconvenient and wished the friend would have a hard time in his/her journey.

We rotated the three levels of the desires manipulation across the scenarios using a Latin Square method (see Table 1 for details). Each participant was presented to either the set of scenarios A or B and went through all levels of the desires

Table 1: Latin Square Design Arrangement

			Set A			
A1	1	Positive	2	Neutral	3	Negative
A2	1	Negative	2	Positive	3	Neutral
A3	1	Neutral	2	Negative	3	Positive
			Set B			
B1	4	Positive	5	Neutral	6	Negative
B2	4	Negative	5	Positive	6	Neutral
B3	4	Neutral	5	Negative	6	Positive

Notes: Numbers "1" to "6" stand for the previously listed scenarios. Each participant went through either the three combinations from set A (i.e., A1, A2, and A3) or the three combinations from set B (i.e., B1, B2, and B3).

manipulation. The presentation of the scenarios was randomly counterbalanced. Hence, we implemented a 2 (Set: A vs. B) × 3 (Desires: positive vs. neutral vs. negative) × 3 (Scenarios) fully crossed, within-participant factorial design, and with Set as a between-subjects variable.

After reading each scenario in the online platform, participants answered on 7-point scales:

- Valence of desires:
  The first item of desires read, "In the scenario, it is informed [the agent]'s desires and thoughts about helping. To which extent do you think [the agent]'s desires and thoughts were negative/positive?" (1 = very negative, 7 = very positive). Participants also rated the extent to which they thought the agent's desires and thoughts were desirable (1 = very undesirable, 7 = very desirable), $rs > .72$, $p < .001$.
- Blame:
  Participants indicated how much they blamed the agent for failing to help (1 = no blame at all, 7 = extreme blame) and for the negative outcome that followed their failure to help (1 = no blame at all, 7 = extreme blame), $rs > 0.76$, $p < .001$.
- Wrongness:
  Participants indicated the extent to which the agent's failure to keep the promise was wrong, (1 = not wrong at all, 7 = extremely wrong).
- Immoral character:
  Participants rated how bad-good, immoral-moral, untrustworthy-trustworthy the character was (1 = extremely moral/ good/ trustworthy, 7 = extremely immoral/ bad/ untrustworthy). We combined these items in our analyses of immoral character, $αs > .84$.
- Causality:
  Participants rated the extent to which the agent was the cause of the negative outcome (1 = not the cause at all, 7 = definitely the cause).
- Intentionality:
  Participants rated the extent to which the agent's failure to help was intentional (1 = not intentional at all, 7 = definitely intentional).

3. Results

Due to our Latin Square design, it was only possible to analyze the desires manipulation as a between-subjects variable when focusing on each individual scenario. Analyses on each scenario for all measures suggested a consistent pattern of results for the main dependent variable, i.e., blame. Therefore, we reanalyzed the data with the desires manipulation as a within-subjects measure, hence combining all scenarios. We conducted One-way repeated measures Analyses of Variance (ANOVAs) on all of the dependent variables. Table 2 shows the means and standard deviations for all measures.

Table 2: Means and standard deviations for each dependent variable

DVs	Valence of desires		
	Positive	Neutral	Negative
Valence of desires	5.09 (1.55) [a]	4.13 (1.27) [b]	3.42 (1.21) [c]
Blame	3.95 (1.58) [b]	4.32 (1.51) [a]	4.28 (1.58) [a]
Wrongness	4.03 (1.63) [a]	4.25 (1.60) [a]	4.33 (1.66) [a]
Immoral character	3.75 (1.26) [c]	4.38 (1.10) [b]	4.75 (1.04) [a]
Causality	4.28 (1.85) [a]	4.30 (1.81) [a]	4.22 (1.86) [a]
Intentionality	2.00 (1.32) [c]	2.24 (1.28) [b]	3.03 (1.55) [a]

Notes: The subscripts "a", "b", and "c" represent the pairwise differences, with higher means ordered in alphabet order. Cells sharing different subscripts in each row were significantly different from each other. Cells sharing the same subscripts were not significantly different from each other.

- Manipulation check:

There was a significant main effect of desires on the valence of desires scores ($F_{(2, 208)} = 104.69$, $p < .001$, $\eta_p^2 = 0.34$), revealing the successfulness of our manipulation. Pairwise comparisons with Bonferroni correction revealed that participants rated the desires of the agent as more positive for the positive desires condition compared to the neutral ($p < .001$) and negative conditions ($p < .001$). The agents' desires in the neutral condition also received greater ratings compared to the negative desires condition ($p < .001$).

- Blame:

Consistent with our predictions, the more the desires of the agent were negative, the greater were the blame scores ($F_{(2, 416)} = 5.12$, $p = .006$, $\eta_p^2 = 0.02$). Specifically, an agent with negative desires was more blameworthy than one with positive desires ($p = .03$). However, agents with negative and neutral desires received blame to a similar extent ($p = 1.00$). This result suggests that it does not take overt negative mental states to invite blame. Displaying indifference to helping others (i.e., neutral condition) was also judged to be equally blameworthy. Finally, agents with positive desires were less blamed than those with neutral desires ($p = .005$).

- Wrongness:

The effect of desires on wrongness scores was only marginally significant ($F_{(2, 416)} = 2.59$, $p = .08$, $\eta_p^2 = 0.01$). There was a tendency for harsher judgments of wrongness the more the desires were negative. However, at a $p < .05$ level of significance, the effect of desires on wrongness was not significant.

- Immoral character:

The more the desires were negative, the more the agent was perceived to be an immoral character ($F_{(2, 416)} = 55.14$, $p < .001$, $\eta_p^2 = 0.21$). An agent with negative desires was evaluated to be more immoral than an agent with positive and neutral desires ($ps < .001$). The agent with neutral desires was also perceived as having poorer character compared to

the one with positive desires ($p < .001$).

- Causality:

The effect of desires on the causality scores was non-significant, $F_{(2, 416)} = 0.09$, $p < .91$), suggesting that ascription of causality was the same for all levels of the desires manipulation.

- Intentionality:

Although all scenarios described the agent's failure to help as unintentional, there was a significant main effect of desires on intentionality, $F_{(2, 416)} = 46.98$, $p < .001$, $\eta_p^2 = 0.19$). The more the desires were negative, the more the failure to help was perceived to be intentional. All pairs were significantly different from each other ($ps < .05$).

- Mediation:

To investigate the psychological processes underlying the effect of desires on blame, we conducted parallel multiple mediation analyses. We performed Bootstrapping analyses (10,000 resampling) using the MEMORE macro for SPSS (Montoya & Hayes, 2017), which is an adequate tool for our repeated measures design. As the effect of the independent variable is calculated by the subtraction of the two repeated-measures observations, we compared the positive versus negative desires conditions. We chose to represent the negativity of the desires by contrasting it with the positive condition (rather than the neutral one) because our previous analyses showed that individuals blamed neutral and negative desires similarly, whereas individuals with negative (vs. positive) desires were deemed to be more blameworthy.

The mediation analyses revealed that the negativity of the desires affected blame indirectly by increasing perceptions of wrongfulness ($ab = 0.09$, 95 % CI = 0.01 to 0.18) and immoral character ($ab = 0.29$, 95 % CI = 0.15 to 0.44). The indirect effects by causality ($ab = -0.01$, 95 % CI = -0.13 to 0.09) and intentionality ($ab = -0.01$, 95 % CI = -0.11 to 0.09) were non-significant. The direct effect was $c = 0.03$, $p = .70$, 95 % CI = -0.23 to 0.15. The total effect was $c = 0.33$, $p < .001$, 95 % CI = 0.32 to 0.35 (see Figure 1).

4. Discussion

In this study, agents who unintentionally failed to help another person were considered to be more blameworthy when their desires were negative (vs. positive). Curiously, our findings also revealed that agents with neutral desires, that is, who did not particularly desire to help but did not oppose it either, received similar blame judgments to those who openly carried negative desires. As discussed in the introduction, this supports the premise that individuals expect others to not only "not have" bad desires but to actually display positive ones.

The significant mediation by wrongness in the effect of negative desires on blame demonstrates a normative aspect of blaming failure to help. Individuals are not expected to help
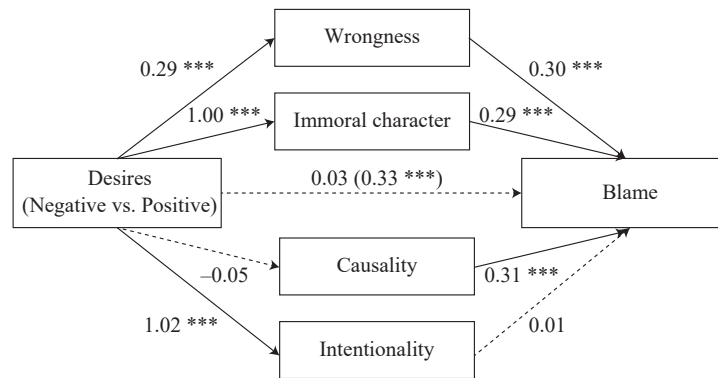
Figure 1: Parallel mediation analyses for the effect of desires on blame for failure to help

Notes: Parallel mediation analyses revealed that the effect of desires (negative vs. positive) on blame for failure to help was mediated by evaluations of wrongness and immoral character. Causality and intentionality were not significant mediators. *** $p < .001$, * $p < .05$.

with negative desires and doing so increases the assignment of blame in case one fails to help. These findings suggest that negative, wicked desires increase blame in part because they are perceived to be wrong. We note that this role of wrongness should be sensitive to individual and cultural differences. For instance, Protestants (vs. Jews) endorse to a greater extent in beliefs about the moral relevance of thoughts, which is associated with a greater role of mental states in person evaluation for this group (Cohen & Rozin, 2001). In cultures which avoid mental state reasoning, mental states are less relevant for moral judgments (McNamara et al., 2019).

The mediation analyses also revealed that an agent who failed to help with negative (vs. positive) desires was deemed to be a more immoral character, which then predicted greater blame. This result cast further light on the apparent illogical role of desires on blame. To the extent that the effect of desires did not affect blame through perceptions of causality and intentionality, rationalist theories of blame would predict that this information would not be relevant in determining blame (see Malle et al., 2014). However, according to the person-centered approach to moral judgments, blame is assigned based on inferences of character underlying one's actions and, in this case, one's desires. Our findings corroborate with this theory and replicate Inbar et al.'s (2012) results with a direct mediation test.

Finally, consistent with the literature on blame, causality judgments predicted blame. However, increased negativity of the desires did not predict inflated judgments of causality, leading to greater blame (i.e., the mediation by causality was nonsignificant). The literature on motivated reasoning suggests that individuals who are motivated to blame may inflate factual moral judgments, such as causal attribution, to support such motivation (Alicke, 1992; Alicke, 2000). In this study, we did not find evidence of this process. Likewise, more negative desires were perceived to be more intentional, consistent with the premise that desires are a fundamental component of intentionality (Malle & Knobe, 1997). Nevertheless, intentionality did not explain the

effect of desires on blame. Overall, these findings rule out the alternative explanation that the mediational roles of wrongness and moral character were a byproduct of underlying perceptions of intentionality and cause.

An important limitation of this study is that we focused our scenarios only on ordinary examples of unintentional failure to help. Perhaps because of the subtle level of harm in these scenarios, the effect of desires on blame was small. An open question is whether the same processes are replicated in judgments of criminal cases, such as omissions. For instance, this study could be extended to cases of medical negligence (e.g., a doctor that has no desire to treat a patient and unintentionally fails to treat them) or parental negligence (e.g., parents who do not desire to care for their children and unintentionally fail to take care of them). Exploring these possibilities may show clearer evidence of the blame processes found in this study. As previously discussed, our findings may also be highly sensitive to cultural influence. Future studies may examine whether these findings will replicate across cultures.

To our knowledge, this is the first study to investigate the role of desires on blame in the context of unintentional failure to help. This study has important implications for the research on moral judgments and justice. We revealed that, in laypersons' judgments, desires determine blame for failure to help to the extent that they signal deviation from the norm and inform immoral character. However, from a legal perspective, desires that do not inform intentionality should be irrelevant for judging blame, and individuals should be punished on the basis of their actions and not their character (Dressler, 2015). This suggests potential biases in the process of judging cases of omission. Further studies may directly explore the processes of blame for failure to help in cases of omission.

Acknowledgments

Note

(1)　The Japanese language does not require gender-specific pronouns, therefore the actual Japanese scenarios did not determine the gender of the agent.

References

Alicke, M. (1992). Culpable causation. *Journal of Personality and Social Psychology*, 63 (3), 368-378.

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126, 556-574.

Ames, D. & Fiske, S. T. (2013). Intentional harms are worse, even when they're not. *Psychological Science*, 24 (9), 1755-1762.

Bostyn, D. H. & Roets, A. (2016). The morality of action: The asymmetry between judgments of praise and blame in the action-omission effect. *Journal of Experimental Social Psychology*, 63, 19-25.

Cohen, A. B. & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, 81 (4), 697-710.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108 (2), 353-380.

Dressler, J. (2015). *Understanding criminal law, 7th ed.* LexisNexis.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Chapter Two–Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55-130.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23 (2), 101-124.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111 (3), 364-371.

Hayashi, H. (2015). Omission bias and perceived intention in children and adults. *British Journal of Developmental Psychology*, 33, 237-251.

Hirozawa, P. Y., Karasawa, M., & Matsuo, A. (2020). Intention matters to make you (im)moral: Positive-negative asymmetry in moral character evaluations. *Journal of Social Psychology*, 160 (4), 401-415.

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 (1), 52-62.

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: Two faces of moral regulation. *Journal of Personality and Social Psychology*, 96 (3), 521-537.

Lagnado, D. & Channon, S. (2008). Jugments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, 108 (3), 754-770.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, 25 (2), 147-186.

Malle, B. F. & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, 33 (2), 101-121.

McNamara, R. A., Willard, A. K., Norenzayan, A., & Henrich, J. (2019). Weighing outcome vs. intent across societies: How cultural models of mind shape moral reasoning. *Cognition*, 182, 95-108.

Montoya, A. & Hayes, A. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, 22 (1), 6-27.

Nadelhoffer, T. (2006). Desire, foresight, intentions, and intentional actions: Probing folk intuitions. *Journal of Cognition and Culture*, 6 (1-2), 133-157.

Pizarro, D. A. & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In M. Mikulincee & P. R. Shaver (eds.), *The social psychology of morality: Exploring the causes of good and evil*, pp. 91-108. APA Press.

Pizarro, D. A., Tannenbaum, D., & Uhlmann, E. (2012). Mindless, harmless, and blameworthy. *Psychological Inquiry*, 23 (2), 185-188.

Pizarro, D. A., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science*, 14 (3), 267-272.

Ritov, I. & Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioral Decision Making*, 3 (4), 263-277.

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. Springer Verlag.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In A. M. Brandt & P. Rozin (eds.), *Morality and health*, p. 119-169. Taylor & Frances/Routledge.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27 (1), pp. 76-105.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, 10 (1), 72-81.

Uhlmann, E., Zhu, L., & Diermeier, D. (2013). When actions speak volumes: The role of inferences about moral character in outrage over racial bigotry. *European Journal of Social Psychology*, 44 (1), 23-29.

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 (2), 283-301.

Appendix

Scenario 1

Ito's best friend was running a donation project for the homeless. The project was very precious to Ito's friend, and Ito's friend would be very grateful if Ito could contribute.

- Positive desires:
  Ito really wanted to donate. (S)he greatly empathized with homeless people and wished the project would be successful and useful to many.
- Neutral desires:
  Ito didn't particularly want to donate but didn't mind donating either. (S)he didn't like nor dislike the project and was overall neutral towards it.
- Negative desires:
  Ito really did not want to donate. S(h)e felt disgusted by homeless people and wished the project would not be successful so that homeless people would learn to make a living on their own.

On the day of the donation, Ito arrived late, just as the donation stall was about to close. Ito ran to the stall and quickly took his/her wallet to donate. However, as (s)he opened it, Ito realized (s)he had spent all the money on groceries. Hence, despite Ito's promise, (s)he ended up not donating.

Days later, Ito found out Ito's best friend had to cancel the project due to lack of donations. "If only I had gotten one more donation, the project would have kept going", the friend told Ito with a sad voice.

Scenario 2

Nakamura and his/her sister were planning their friend's birthday celebration. The sister wanted to do something special, as their friend had been very sad lately. Then, the sister had the idea of surprising their friend with a very fancy birthday cake after dinner. They knew their friend had always wanted to try that cake but would not buy it as it was too expensive. As the sister would cook the dinner, Nakamura's sister asked Nakamura to buy the cake, and they would split the expenses.

- Positive desires:
  Nakamura really wanted to buy the cake. Nakamura loved celebrations and wished the friend would be happy with the surprise.
- Neutral desires:
  Nakamura didn't particularly want to buy the cake but didn't mind buying it either. (S)he didn't like nor dislike celebrations and was overall neutral about it.
- Negative desires:
  Nakamura really did not want to buy the cake. (S)he hated celebrations and wished the friend would overcome their

emotions by themselves.

[However], According to the plan, Nakamura promised to buy the cake.

On the birthday night, Nakamura went to the cake shop after work, but it was closed. (S)he had made a mistake; Nakamura thought the store would be open until 7pm, but it closed at 6pm.

In the end, as they had no surprise cake, the dinner went by flat and did not feel like a celebration at all. Both friend and sister were very disappointed but didn't say anything.

Scenario 3

Sato's cousin was moving out from their apartment, but the cousin didn't have a car. Sato had a very big car, so his/her mother told him/her to help the cousin.

- Positive desires:
  Sato really wanted to help the cousin. (S)he was happy to be useful for others and wished the cousin would have a smooth moving out process.
- Neutral desires:
  Sato didn't particularly want to help the cousin but didn't mind it either. Sato didn't like nor dislike cooperating with others and had no special opinion about it.
- Negative desires:
  Sato really did not want to help the cousin. (S)he was annoyed to be bothered by others and for that Sato wished the cousin would have a stressful moving out process.

In the end, in line with his/her mother's request, Sato promised the cousin to help with the moving process. The cousin showed great gratitude.

On the moving out day, at the settled time, Sato was about to go to the cousin's apartment, when Sato remembered there was an important meeting at work. Sato rushed to work and ended up cancelling on the cousin.

At night, as Sato was returning home, Sato received another call. His/her mother said the cousin was not able to finish moving out by himself/herself. Hence, the cousin had to pay an expensive fee to the real state agency, as there was still some furniture in the apartment on the day (s)he should have returned the apartment.

Scenario 4

Yamada's classmate was having lots of difficulty with math. Yamada was very good at math, so Yamada's teacher said it would be nice if Yamada could support the classmate.

- Positive desires:
  Yamada really wanted to help the classmate. (S)he thought that people should always help each other and wished the classmate would persevere on his/her studies.

- Neutral desires:
Yamada didn't particularly want to help the classmate but didn't mind helping either. (S)he didn't like nor dislike co-operating with others and had no special opinion about it.
- Negative desires:
Yamada really did not want to help the classmate. (S)he thought people should not be inconvenient to others and wished the classmate would just give up on his/her studies.

In the end, as suggested by his/her teacher, Yamada promised to help the classmate with math for the upcoming math test. They set up they would meet on Sunday morning to study, as the test was on Monday. The classmate showed great appreciation.

On Sunday, Yamada spent the day doing groceries and cleaning the house with his/her mother. At night, when (s)he decided to review for the test, (s)he realized (s)he had forgotten about the promise to the classmate.

On the next day, they had the math test. A week later, Yamada heard the classmate had failed the test.

Scenario 5

Hayashi lived alone in his/her apartment. Hayashi's mother lived with his/her sick grandfather at the other side of the city. One day, Hayashi's mother called. She said she had to do a night shift, so she asked if Hayashi could spend the night with the grandfather and make sure he would take his medicine.

- Positive desires:
Hayashi really wanted to take care of the grandfather. (S)he liked him since childhood and wished he would recover soon.
- Neutral desires:
Hayashi didn't specifically want to take care of the grandfather but didn't mind it either. (S)he had a distant relationship with the grandfather and was overall neutral towards him.
- Negative desires:
Hayashi really did not want to take care of the grandfather. (S)he disliked him since childhood and wished he wouldn't recover soon.

[However], in the end, Hayashi promised to take care of the grandfather.

That day, Hayashi was exhausted after work and fell asleep immediately after getting to bed. The next day, Hayashi woke up realizing (s)he had forgotten about the promise to take care of the grandfather.

Later, (s)he found out that, as the grandfather had not taken the medication properly the night before, his condition had worsened.

Scenario 6

Hirata's friend was traveling abroad. His/her flight was scheduled to depart at 5am. As Hirata had a car, Hirata's friend asked Hirata to drop him/her off at the airport around 3am, since public services did not work at that time.

- Positive desires:
Hirata really wanted to drive the friend to the airport. (S)he thought friends should help each other and wished the friend would have a pleasant journey.
- Neutral desires:
Hirata didn't particularly want to drive the friend to the airport but didn't mind doing it either. (S)he felt neutral about the friend's request.
- Negative desires:
Hirata really did not want to drive the friend to the airport. (S)he thought the friend was inconvenient and wished the friend would have a hard time in his/her journey.

[However], in the end, Hirata promised to drop the friend off at the scheduled time.

On the night before the flight, Hirata was suddenly burdened with lot of work at the office. (S)he worked until late and fell asleep on the sofa after returning home. The next morning, (s)he realized (s)he had forgotten about driving the friend to the airport.

Hirata later found out the friend had lost his/her flight.