

# Journal of Experimental Psychology: General

## Reasons Probably Won't Change Your Mind: The Role of Reasons in Revising Moral Decisions

Matthew L. Stanley, Ashley M. Dougherty, Brenda W. Yang, Paul Henne, and Felipe De Brigard  
Online First Publication, September 18, 2017. <http://dx.doi.org/10.1037/xge0000368>

### CITATION

Stanley, M. L., Dougherty, A. M., Yang, B. W., Henne, P., & De Brigard, F. (2017, September 18). Reasons Probably Won't Change Your Mind: The Role of Reasons in Revising Moral Decisions. *Journal of Experimental Psychology: General*. Advance online publication. <http://dx.doi.org/10.1037/xge0000368>

# Reasons Probably Won't Change Your Mind: The Role of Reasons in Revising Moral Decisions

Matthew L. Stanley  
Duke University

Ashley M. Dougherty  
Wake Forest School of Medicine, Winston-Salem,  
North Carolina

Brenda W. Yang, Paul Henne, and Felipe De Brigard  
Duke University

Although many philosophers argue that making and revising moral decisions ought to be a matter of deliberating over reasons, the extent to which the consideration of reasons informs people's moral decisions and prompts them to change their decisions remains unclear. Here, after making an initial decision in 2-option moral dilemmas, participants examined reasons for only the option initially chosen (affirming reasons), reasons for only the option not initially chosen (opposing reasons), or reasons for both options. Although participants were more likely to change their initial decisions when presented with only opposing reasons compared with only affirming reasons, these effect sizes were consistently small. After evaluating reasons, participants were significantly more likely not to change their initial decisions than to change them, regardless of the set of reasons they considered. The initial decision accounted for most of the variance in predicting the final decision, whereas the reasons evaluated accounted for a relatively small proportion of the variance in predicting the final decision. This resistance to changing moral decisions is at least partly attributable to a biased, motivated evaluation of the available reasons: participants rated the reasons supporting their initial decisions more favorably than the reasons opposing their initial decisions, regardless of the reported strategy used to make the initial decision. Overall, our results suggest that the consideration of reasons rarely induces people to change their initial decisions in moral dilemmas.

*Keywords:* affect, choice, emotion, moral, reasoning

*Supplemental materials:* <http://dx.doi.org/10.1037/xge0000368.supp>

When faced with a moral decision, there are often many, diverse reasons for and against the possible choices. Many philosophers argue that making and revising moral decisions ought to be a matter of deliberating over reasons, because reasons serve to favor or to justify choices for action (Broome, 2013; Crisp, 2005; Liberman & Schroeder, 2016; Parfit, 2011; Philips, 1987; Ross, 1930; Scanlon, 1998; Snedegar, 2017). In this study, we address several descriptive questions about the role of reasons in making and

revising moral decisions. Specifically, we investigate the extent to which reasons inform moral decisions, the extent to which people revise their initial decisions after considering reasons, and the extent to which people engage in a biased, motivated evaluation of reasons to support their initial decisions.

Despite the common prescriptive claim that we ought to use reasons to make and revise moral decisions, accumulating empirical evidence suggests that conscious deliberation and reasoning

---

Matthew L. Stanley, Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University; Ashley M. Dougherty, Wake Forest School of Medicine, Winston-Salem, North Carolina; Brenda W. Yang, Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University; Paul Henne, Department of Philosophy, Duke University; Felipe De Brigard, Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Department of Philosophy, Duke Institute for Brain Sciences, Duke University.

The Imagination and Modal Cognition (IMC) Lab at Duke University is broadly interested in issues that sit at the intersection between philosophy and the cognitive sciences. We believe that moral decision-making can be usefully informed through both philosophical and empirical work. The present paper on how reasons influence moral decisions in philosophical as well as everyday moral dilemmas intersects with our past work on how

cognitive psychology and neuroscience can inform the judgments people make about memories of immoral actions as well as judgments of blame and responsibility. We believe the results in this article showing how ordinary people approach moral decisions is a useful contribution to both literatures. Moving forward, we find the general resistance to changing decisions in the face of reasons a fascinating phenomenon that deserves more attention, one worth studying in a world where evidence and information change rapidly.

We thank members of the IMC Lab and the MAD Lab at Duke University for their helpful comments and suggestions. This project was supported by a research grant from the Duke Institute for Brain Science to FDB.

Correspondence concerning this article should be addressed to Felipe De Brigard, Department of Philosophy, Duke University, 203A West Duke Building, Durham, NC 27708-0743. E-mail: [felipe.debrigard@duke.edu](mailto:felipe.debrigard@duke.edu)

rarely contribute to the initial moral judgments and decisions that people make in private. Following Hume (1777/1965) and Zajonc (1980), Haidt (2001, 2007) argued that moral judgments and decisions are predominantly driven by automatic, immediate responses that are effortless, reactive, affective and produced by unconscious processes. Consistent with this theory, people typically have immediate implicit reactions to scenes or stories of moral violations (Luo et al., 2006). Immediate affective reactions are usually good predictors of moral judgments and behaviors (Haidt & Hersh, 2001; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), and affective manipulations can influence moral judgments (Ugazio, Lamm, & Singer, 2012; Valdesolo & DeSteno, 2006; Wheatley & Haidt, 2005). Notably, research on moral dumbfounding suggests that people insist that seemingly harmless taboo violations are morally wrong in the face of conflicting reasons—even when they cannot explain why the behavior is morally wrong (Haidt, Bjorklund, & Murphy, 2000; Haidt, 2001; although, see Royzman, Kim, & Leeman, 2015). Haidt (2001, 2007) has argued that affective, reactive moral judgments and decisions typically come first, and if people engage in reasoning at all, such reasoning is post hoc and typically motivated by a search for and a preferential weighing of evidence supporting initial, affective reactions (Ditto, Pizarro, & Tannenbaum, 2009; Haidt, 2007, 2012; Haidt, Koller, & Dias, 1993).

Nevertheless, reasons, reasoning, and deliberation likely contribute to making and to revising moral judgments and decisions under certain circumstances (Bloom, 2010; Haidt, 2012; Horne, Powell, & Hummel, 2015; Maki & Raimi, 2017; Paxton & Greene, 2010; Paxton, Ungar, & Greene, 2012; Wiech et al., 2013). However, there is relatively little empirical evidence informing the extent to which the consideration of different reasons leads people to change their moral judgments and decisions. There is evidence from other domains of research (e.g., behavioral economics) that attitudes, preferences, and decisions are often made or modified by the consideration and weighing of different reasons (Barber, Heath, & Odean, 2003; Dietrich & List, 2016; Petty & Cacioppo, 1986; Shafir, Simonson, & Tversky, 1993; Simonson, 1989; Vlaev, Chater, Lewis, & Davies, 2009; Wilson et al., 1989). When faced with the need to make or revise a decision, people often seek out reasons and evaluate the quality of reasons to resolve conflict and justify their choices (Shafir et al., 1993). In the moral domain, people may similarly use available reasons to revise their decisions.

Using a novel experimental paradigm, we investigate three descriptive questions about the role of reasons in revising moral decisions. First, how might considering different sets of reasons affect the likelihood people change their decisions? Specifically, are there differences in the likelihood that people change their initial decisions in two-choice moral dilemmas after considering reasons for only the option initially chosen (affirming reasons), reasons for only the option not initially chosen (opposing reasons), or the complete set of reasons for both options? Regardless of whether initial moral judgments and decisions are typically the product of immediate, affective reactions, reasons and reasoning may still play a prominent role in revising moral decisions (Bloom, 2010; Bucciarelli & Daniele, 2015; Bucciarelli, Khemlani, & Johnson-Laird, 2008; Pinillos, Smith, Nair, Marchetto, & Mun, 2011). Even antirationalists (e.g., Haidt, 2001; Greene, 2007; Greene & Haidt, 2002) have suggested that there are several

possible ways to override or alter decisions, all of which involve reasons and reasoning. Furthermore, Pizarro and Bloom (2003), citing historical evidence, suggested that when confronted with information contrary to their moral beliefs, people often do make different decisions after deliberation. If people are, in fact, willing to override their initial decisions after considering conflicting or opposing reasons, then people would be significantly more likely to alter their initial decisions when only presented with reasons supporting the option not initially chosen versus reasons supporting the option initially chosen.

Second, what will be a better predictor of the final decision: (a) the initial decision made before evaluating reasons or (b) the particular set of reasons evaluated (i.e., affirming reasons, opposing reasons, or reasons for both options)? If people are willing to use reasons to arrive at decisions and change their initial decisions when new, relevant reasons are presented, then the particular reasons presented should have some utility in predicting the final decision above and beyond the initial decision. However, if people are resistant to changing their decisions in the face of reasons, then the initial decision should be a better predictor of the final decision than the particular reasons considered. That is, people should typically make the same decision before and after evaluating reasons.

Third, do people engage in a biased, motivated evaluation of available reasons to support their initial, affective reactions? If people rarely change their initial decisions after considering different reasons, then this resistance to changing their decisions could be explained by motivated evaluation of the reasons presented. After making an initial decision, people may evaluate reasons in support of that decision relatively favorably or evaluate reasons opposing that decision relatively unfavorably. Then, after concluding that the reasons supporting the initial decision are relatively good reasons or that the reasons opposing the initial decision are relatively poor reasons through this biased evaluation process, people would be more likely to make the exact same decision. Such a finding would suggest that people evaluate reasons in a biased, motivated way to support their initial, affective decisions in the moral domain.

Two sets of experiments follow to investigate the questions above. Studies 1a, 1b, and 1c examine the influence of reasons in moral decision-making using variants of classic moral dilemmas. Studies 2a, 2b, and 2c utilize more commonplace, everyday dilemmas to investigate the role of reasons in moral decision-making.

## Study 1a

The purpose of Study 1a is to norm and validate a set of classic two-option moral dilemmas and accompanying reasons for making each possible decision.

## Method

**Participants.** 256 individuals voluntarily participated in this study via Amazon's Mechanical Turk (AMT) for monetary compensation. For reliability of AMT data for experimental studies, see Rand (2012). Participant recruitment was restricted to individuals in the United States with a prior approval rating above 80%. Thirteen participants failed to follow instructions; thus, data were

analyzed with the remaining 243 individuals ( $M_{\text{age}} = 33.16$ ,  $SD = 9.21$ ,  $\text{range}_{\text{age}} = [19-65]$ , 100 females). Racial/ethnic groups included 80.0% White non-Hispanic, 10.7% Black non-Hispanic, 6.6% Hispanic, 2.5% Asian non-Hispanic, .4% Native American, and 2.5% Other. Overall, participants in the sample were well-educated (.4% did not complete high school, 16.9% only completed high school, 39.5% completed some college/associate degree, 38.3% completed a bachelor's degree, 4.5% completed a master's degree, and .4% completed a PhD). This study was approved by the Duke University Campus Institutional Review Board.

**Materials.** Three moral dilemmas were used. Each dilemma was adapted from published research (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008) and altered to facilitate conflict between reasons for competing possible choices. In each dilemma, there were only two possible decisions that participants could make. A separate group of 20 participants recruited from AMT participated in a norming session. These 20 participants were excluded from subsequent studies. The three moral scenarios were presented separately to each participant in a randomized order. Participants were asked to provide their best four reasons for choosing each option in each moral scenario presented herein. Our goal was to generate reasons that were compelling and representative of the reasons that the typical person might consider when contemplating moral dilemmas. We compiled the six most commonly reported reasons for choosing each option, modified for clarity and brevity. These reasons were diverse: whereas some captured distinct moral principles or obligations (e.g., "I would be saving a greater number of lives"), others appealed to anticipated emotional reactions (e.g., "I would have to live with the guilt for the rest of my life") or to the possibility of self-preservation (e.g., "Choosing this option is the only way that I would survive"). See Appendix A for the moral scenarios and reasons used in Study 1a-c.

**Procedure.** The experimental procedure is depicted in Figure 1a. Participants first read a randomly selected moral dilemma and imagined themselves in the scenario. They then made judgments about six reasons for choosing each of the two possible options,

thus rating a total of 12 reasons. Specifically, participants were asked to evaluate the quality of each reason for choosing that particular option in the scenario on a scale from  $-10$  (*worst possible reason*) to  $10$  (*best possible reason*). All reasons were presented on the same screen, but reasons for each possible choice were grouped together. After making all ratings for a given dilemma, participants were asked to indicate which option they would choose in the scenario. Finally, participants were asked if they considered any additional reasons not presented in the study when making their decision. This same procedure was then repeated for a second randomly selected moral dilemma and then for the third and final moral dilemma. The entire study was self-paced. The order of the moral dilemmas presented as well as the order of the reasons for each option were randomized without replacement.

## Results and Discussion

In sum, 42.39% of participants indicated that they would choose to smother the baby in the crying baby dilemma, 40.33% of participants indicated that they would choose to pull the chair out from underneath the son in the concentration camp dilemma, and 20.58% of participants indicated that they would choose to transplant the organs in the organ transplant dilemma. Furthermore, only 7.41%, 7.00%, and 7.82% of participants reported having considered an additional reason or additional reasons beyond those provided in the crying baby, concentration camp, and organ transplant dilemmas, respectively. This finding suggests that our normed set of reasons for choosing each option in the moral dilemmas is relatively exhaustive. Supplementary Figure 1 depicts distributions of ratings for each reason in each moral dilemma across participants.

### Study 1b

In Study 1b, we consider the influence of reasons on changing decisions in moral dilemmas. After participants make an initial choice in a moral dilemma, they were presented with reasons congruent with their initial choice (affirming reasons), reasons

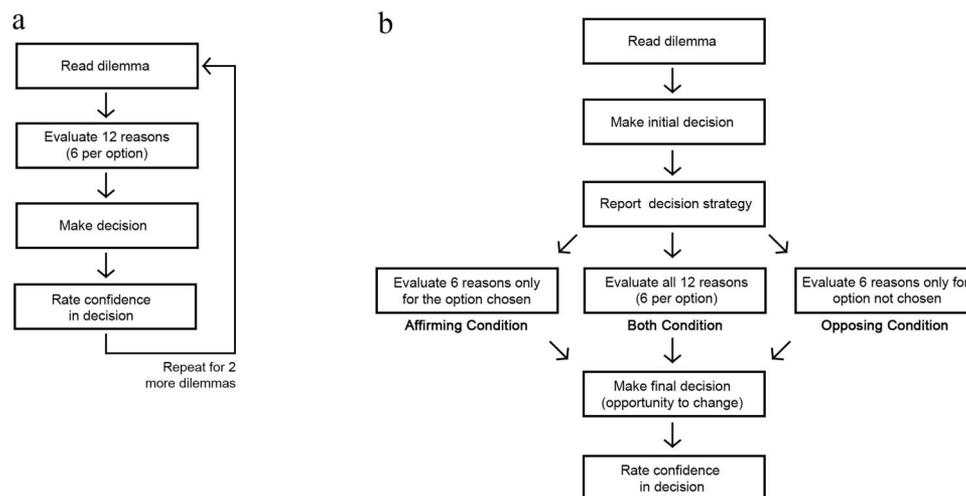


Figure 1. Schematic of the experimental paradigms for (a) Study 1a and (b) Study 1b.

supporting the opposing option (opposing reasons), or the full set of reasons (reasons for both options). We investigate whether participants are more likely to change their decisions after considering opposing reasons or affirming reasons. We also investigate whether the initial decision is a better predictor of the final decision than the particular set of reasons presented. Furthermore, we investigate whether people engage in a biased, motivated evaluation of reasons to support their initial, affective reactions.

## Method

**Participants.** Two thousand, two hundred sixty individuals voluntarily participated in this study via AMT for monetary compensation. Participant recruitment was restricted to individuals in the United States with a prior approval rating above 80%. Forty-five participants failed to follow instructions, so data were analyzed with the remaining 2,215 individuals ( $M_{\text{age}} = 35.11$ ,  $SD = 11.31$ ,  $\text{range}_{\text{age}} = [18-84]$ , 976 females). This study was approved by the Duke University Campus Institutional Review Board.

**Materials and procedure.** The set of moral dilemmas and reasons from Study 1a was used in Study 1b (see Appendix A for details). The procedure is depicted in Figure 1b. The entire study was self-paced. Participants were instructed to carefully read a randomly selected moral dilemma. Participants were randomly assigned to exactly one of the three moral dilemmas. After reading the dilemma, participants were instructed to make an initial decision indicating which option they would choose in the dilemma and to indicate how they made the decision. Participants were instructed to indicate whether they based their decision on (a) an immediate reaction or feeling after reading the dilemma such that no reasons affected the decision, (b) at least one reason, or (c) some other method. If a participant selected option (b), he or she indicated what reason or reasons brought about the decision. If a participant selected option (c), he or she indicated how the decision was made.

Then, participants were randomly assigned to exactly one of three between-subjects conditions: only affirming reasons were considered, only opposing reasons were considered, or all reasons for both options were considered. The order in which the reasons were presented was randomized. Participants were asked to rate how compelling each

reason was for choosing that particular option in the dilemma on a scale from  $-10$  (*worst possible reason*) to  $10$  (*best possible reason*). After providing ratings for reasons in a dilemma, participants were asked to indicate which option they would choose. Then, participants rated how confident they were in the final decision ( $-5 = \text{not at all confident}$ ,  $5 = \text{very confident}$ ). Finally, participants were asked if they considered any additional reasons not presented in the study. If participants considered other reasons, they were instructed to describe those additional reasons. Upon completion, participants were monetarily compensated for their time.

## Results

Table 1 shows the percentage of participants indicating they would make each possible decision in each dilemma as a function of the set of reasons presented and when the decision was made (the initial decision before seeing reasons or the final decision after seeing reasons). Furthermore, 53.10%, 51.28%, and 57.83% of participants reported having made their initial decisions based upon at least one reason in the crying baby, concentration camp, and organ transplant dilemmas, respectively; 45.38%, 46.97%, and 40.43% of participants reported having made their initial decisions based upon an immediate reaction or feeling in the crying baby, concentration camp, and organ transplant dilemmas, respectively. The remaining participants indicated that they made their decisions in some other way.

**The effect of reasons on changing moral decisions.** First, we investigated whether participants were more likely to change their initial decisions after considering affirming reasons, opposing reasons, or reasons for both options. Collapsing across all three dilemmas, there was a small but statistically significant relationship between whether or not participants changed their decisions and the particular reasons evaluated (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .026$ ,  $p < .001$ ). Evaluating the dilemmas separately, there were small but statistically significant relationships between whether or not participants changed their decisions and the particular reasons evaluated in the crying baby (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .028$ ,  $p < .001$ ), concentration camp (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .032$ ,  $p <$

Table 1  
*Percentages of Participants Making Each Possible Decision*

Decision	All reasons	Affirming reasons	Opposing reasons
<b>Initial decision</b>			
Crying baby (smother baby)	40% ( $n = 246$ )	47% ( $n = 236$ )	38% ( $n = 243$ )
Concentration camp (pull chair)	36% ( $n = 248$ )	40% ( $n = 252$ )	36% ( $n = 243$ )
Organ transplant (transplant organs)	13% ( $n = 246$ )	18% ( $n = 250$ )	16% ( $n = 251$ )
<b>Final decision</b>			
Crying baby (smother baby)	38% ( $n = 246$ )	46% ( $n = 236$ )	35% ( $n = 243$ )
Concentration camp (pull chair)	36% ( $n = 248$ )	41% ( $n = 252$ )	36% ( $n = 243$ )
Organ transplant (transplant organs)	11% ( $n = 246$ )	16% ( $n = 250$ )	12% ( $n = 251$ )

*Note.* Percentages of participants who indicated they would choose the option specified in parentheses in each dilemma as a function of condition (affirming reasons, opposing reasons, or reasons for both options) and when the decision was made (the initial decision before seeing reasons or the final decision after seeing reasons) are depicted. Specifically, we report the percentage of participants who indicated that they would choose to smother the crying baby, pull the chair out from underneath the son, and transplant the organs. The total number of participants in each condition is also reported.

.001), and organ transplant (Fisher's exact test:  $p = .002$ , two-sided; Goodman and Kruskal  $\tau = .018$ ,  $p = .001$ ) dilemmas. Figure 2 depicts the percentages of participants who changed their decisions in each dilemma as a function of the reasons considered. Relatively few participants changed their initial decisions regardless of which reasons were considered.

Employing ransacking (Goodman, 1969), we found that participants who considered only opposing reasons compared with only affirming reasons were more likely to change their initial decisions in the crying baby (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .039$ ,  $p < .001$ ), concentration camp (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .048$ ,  $p < .001$ ), and organ transplant (Fisher's exact test:  $p = .001$ , two-sided; Goodman and Kruskal  $\tau = .022$ ,  $p = .001$ ) dilemmas. Similarly, participants who considered only opposing reasons compared with reasons for both options were more likely to change their initial decisions in the crying baby (Fisher's exact test:  $p = .014$ , two-sided; Goodman and Kruskal  $\tau = .013$ ,  $p = .014$ ), concentration camp (Fisher's exact test:  $p = .027$ , two-sided; Goodman and Kruskal  $\tau = .011$ ,  $p = .027$ ), and organ transplant (Fisher's exact test:  $p = .031$ , two-sided; Goodman and Kruskal  $\tau = .010$ ,  $p = .031$ ) dilemmas. Participants who considered reasons for both options were more likely to change their initial decisions than participants who considered only affirming reasons in the crying baby dilemma (Fisher's exact test:  $p = .046$ , two-sided; Goodman and Kruskal  $\tau = .009$ ,  $p = .046$ ) and the concentration camp dilemma (Fisher's exact test:  $p = .004$ , two-sided; Goodman and Kruskal  $\tau = .017$ ,  $p = .004$ ) but not the organ transplant dilemma ( $p > .25$ ). All results are reported prior to correction for multiple comparisons. Across all three dilemmas, while participants were significantly more likely to change their decisions when only presented with opposing reasons compared with being presented with affirming reasons or reasons for both options, all effect sizes were small. Within each condition and dilemma, no more than 12% of participants changed their initial decisions after considering reasons (Figure 2).

Next, we sought to determine the extent to which the initial decision predicts the final decision made after evaluating different reasons. To this end, three separate logistic regression models were computed, one for each moral dilemma. In each model, the dependent variable is the final decision made in the moral dilemma after considering reasons, and the two independent variables are

(1) the option initially chosen in the dilemma before examining reasons and (2) the reasons considered (i.e., condition: affirming reasons, opposing reasons, or reasons for both options). Even after accounting for the particular reasons evaluated, the initial decision was a strong and significant predictor of the final decision in the crying baby ( $b = 5.53$ ,  $SE = .33$ ,  $e^b = 251.57$ ,  $Z = 16.77$ ,  $p < .001$ , 95% CI = [4.91, 6.21]), concentration camp ( $b = 5.34$ ,  $SE = .33$ ,  $e^b = 207.91$ ,  $Z = 17.55$ ,  $p < .001$ , 95% CI = [4.77, 5.96]), and organ transplant ( $b = 5.95$ ,  $SE = .48$ ,  $e^b = 382.56$ ,  $Z = 12.49$ ,  $p < .001$ , 95% CI = [5.09, 6.98]) dilemmas. Participants typically chose the same option before and after evaluating reasons, regardless of which reasons were evaluated. Unsurprisingly, in separate follow-up models including only the subset of participants who considered only opposing reasons, the initial decision was still a strong predictor of the final decision in all dilemmas (all  $ps < .001$ ), indicating that participants rarely changed their decisions from the initial choice to the final choice even in the face of opposing reasons (Figure 2). Additionally, regardless of which reasons were considered, participants tended to report high confidence in their final decisions (supplementary Table 1).

We next sought to compare pseudo- $R^2$  values between models that included versus did not include condition (i.e., affirming reasons, opposing reasons, or reasons for both options) to determine how much variance each condition explains above and beyond the initial decision. After running three additional logistic regression models excluding condition as an independent variable, the amount of variance explained in the models for each dilemma changed minimally compared with the full models with all predictors. The models that only included the initial decision accounted for most of the variance in the final decision made. Table 2 reports pseudo- $R^2$  values for each model.

**Motivated evaluation of reasons.** Having found that participants rarely changed their decisions regardless of the reasons considered, we investigated whether this resistance to changing decisions is at least partly attributable to a biased, motivated evaluation of reasons. We first investigated whether participants engaged in this kind of biased, motivated evaluation when considering reasons for both options in the dilemmas. Specifically, after making an initial decision and then evaluating reasons for both options, we investigated whether participants evaluated reasons supporting that decision more favorably than reasons supporting the other possible decision in the dilemma. To this end, we computed separate two-way ANOVAs for each dilemma. For each ANOVA, the initial decision (binary factor) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable for each ANOVA was computed by first averaging the rated quality of reasons for each possible decision separately and then taking the difference in average ratings for one option relative to the other option.

Our results indicate that participants do evaluate the reasons supporting their initial decisions more favorably than the reasons supporting the foregone option in each moral dilemma, regardless of whether the initial decision was reportedly the product of at least one reason or an immediate, affective reaction. Table 3 displays results from each two-way ANOVA. In each model, there was a large and significant main effect of initial decision, but no effect of initial strategy and no interaction between initial decision and initial strategy. Specifically, the main effects of initial decision

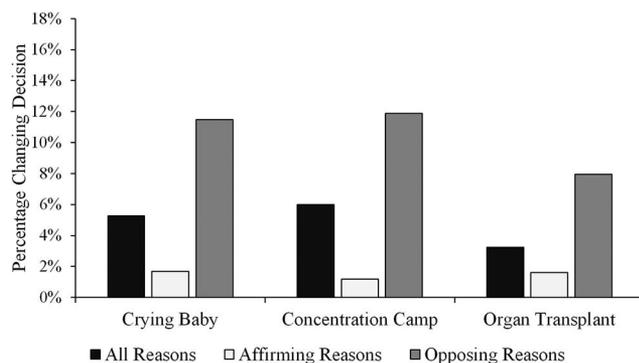


Figure 2. The percentages of participants changing their decisions in each dilemma and condition are depicted.

Table 2  
*Model Comparisons*

Model	Crying baby	Concentration camp	Organ transplant
Model with condition	.802	.786	.751
Model without condition	.800	.785	.747

*Note.* Nagelkerke  $R^2$  values are depicted for each of six different logistic regression models. For each dilemma, one logistic regression model was computed including condition (i.e., a categorical variable with levels for affirming reasons, opposing reasons, or reasons for both options) and the initial decision as predictors of the final decision, while another logistic regression model was run only including the initial decision as a predictor of the final decision.

indicated that individuals rate reasons supporting their initial decisions more favorably than those supporting the foregone option for the crying baby (Mean difference = 7.01,  $SE = .67$ ,  $p < .001$ , 95% CI = [5.68, 8.33]; Figure 3a), concentration camp (Mean difference = 3.79,  $SE = .52$ ,  $p < .001$ , 95% CI = [2.78, 4.81]; Figure 3b), and organ transplant (Mean difference = 6.92,  $SE = .96$ ,  $p < .001$ , 95% CI = [5.02, 8.81]; Figure 3c) dilemmas. There were no significant interactions in any of the models.

Second, we investigated whether participants engaged in a biased, motivated reasoning when examining only affirming reasons versus only opposing reasons. Specifically, we investigated whether individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons, regardless of the reported strategy used to make the initial decision. To this end, separate two-way ANOVAs were computed for each dilemma. For each ANOVA, the condition (binary factor: only opposing reasons evaluated or only affirming reasons evaluated) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable was the average rated quality of the reasons evaluated. Table 4 displays results from each two-way ANOVA. In each model, there was a large and significant main effect of condition. Specifically, the main effects of condition indicated that individuals rated affirming reasons more favorably than opposing reasons for the crying baby (Mean difference = 4.46,  $SE = .39$ ,  $p < .001$ , 95% CI = [3.70, 5.23]; Figure 3d), concentration camp (Mean difference = 1.51,  $SE = .30$ ,  $p < .001$ , 95% CI = [.92, 2.11]; Figure 3e), and organ transplant (Mean difference = 5.69,  $SE = .36$ ,  $p < .001$ , 95% CI = [4.98, 6.39]; Figure 3f) dilemmas. The significant main effect of condition in the organ transplant dilemma was qualified by a small but significant interaction between condition and initial strategy (Table 4), but subsequent tests of simple main effects showed that individuals evaluate affirming reasons more favorably than opposing reasons regardless of the reported strategy used to make the initial decision (both  $ps < .001$ ; Figure 3f).

## Discussion

Across all three dilemmas, our results suggest that people are more likely to change their initial decisions when presented with only opposing reasons compared with only affirming reasons. Nevertheless, the strength of the associations for these significant

results suggest that knowing which set of reasons were considered by participants only reduces prediction error rate by less than 5%. Additional analyses revealed that participants were significantly more likely not to change their initial decisions than to change those decisions after evaluating reasons, regardless of the set of reasons considered. The initial decision before seeing reasons accounted for the vast majority of the variance in predicting the final decision, whereas the particular set of reasons considered only minimally increased goodness of fit (i.e., pseudo- $R^2$  values) when added to the model. These results suggest that participants' initial decisions before seeing reasons are much better predictors of the final decisions than the set of reasons considered. Participants very rarely changed their initial decisions.

Additionally, across all three moral dilemmas, after making an initial decision and then evaluating reasons for both options, participants evaluated reasons supporting the initial decision more favorably than reasons supporting the other possible option. Similarly, individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons. These findings held regardless of whether participants reported making their initial decisions based on at least one reason or an immediate, affective reaction. Thus, individuals do seem to evaluate reasons in a motivated way to support their previous decisions. This behavior is likely one important factor explaining why individuals rarely change their moral decisions in the face of reasons.

## Study 1c

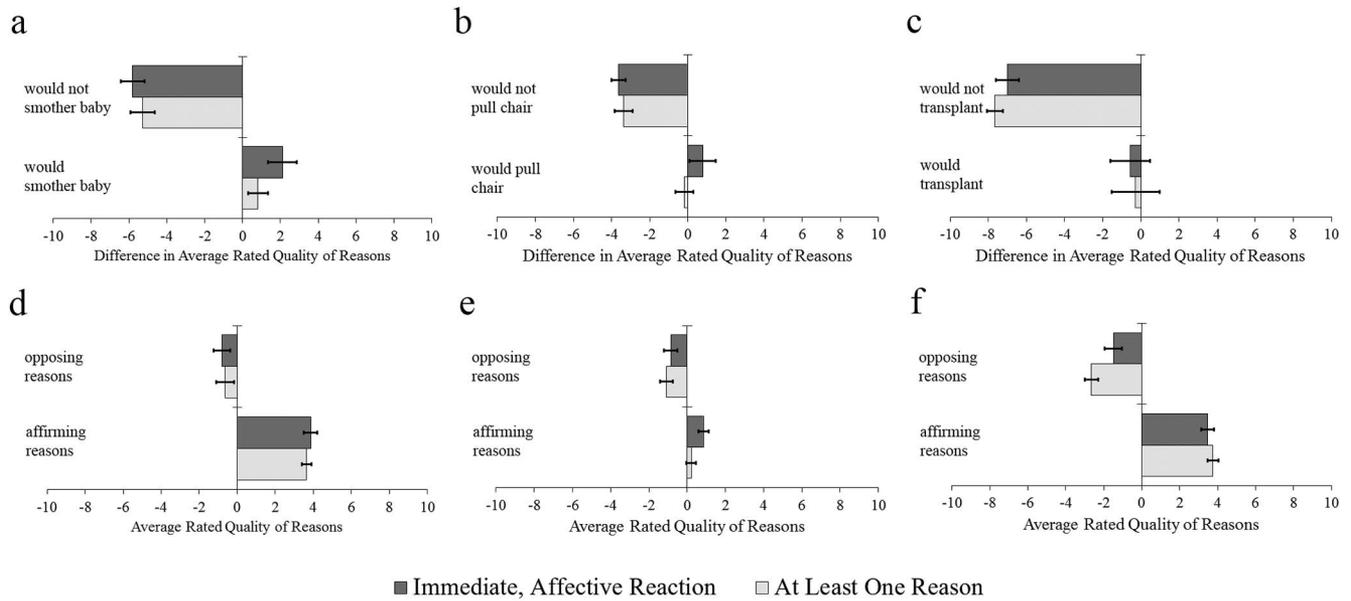
Building upon the results from Study 1b, we further considered the influence of reasons on changing moral decisions in Study 1c. As in Study 1b, we investigate whether participants are more likely to change their initial decisions after considering opposing reasons or affirming reasons, whether the initial decision is a better predictor of the final decision than the particular set of reasons presented, and whether participants engage in a biased, motivated evaluation of reasons to support their initial, affective reactions.

Critically, Study 1c extends Study 1b in three key ways. First, it is possible that changing one's mind requires time (Horne et al.,

Table 3  
*Full Results From ANOVAs*

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
Crying baby				
Initial decision	1	108.96	<.001	.313
Initial strategy	1	.34	.560	.001
Initial Decision $\times$ Initial strategy	1	1.83	.178	.008
Concentration camp				
Initial decision	1	54.36	<.001	.186
Initial strategy	1	.45	.503	.002
Initial Decision $\times$ Initial strategy	1	1.45	.229	.006
Organ transplant				
Initial decision	1	51.79	<.001	.179
Initial strategy	1	.04	.842	.000
Initial Decision $\times$ Initial strategy	1	.23	.634	.001

*Note.* Full results from the three separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including the initial decision and the strategy used to make that initial decision as independent variables. The dependent variable was the difference in the average rated quality of reasons between the two possible choices.



**Figure 3.** For those participants who evaluated reasons for both options, means and standard error bars are shown for differences in the average rated quality of reasons between the two possible choices in the crying baby (a), concentration camp (b), and organ transplant (c) dilemmas, split by the initial decision and the reported strategy used to make the initial decision. Difference scores were computed by subtracting the average rating of reasons for not committing the action from the average rating of reasons for committing the action. For example, the difference in average ratings for the crying baby dilemma was computed by subtracting the average rating for not smothering the baby from the average rating for smothering the baby. For participants who evaluated only affirming reasons or only opposing reasons, means and standard error bars are depicted for the average rated quality of the reasons evaluated in the crying baby (d), concentration camp (e), and organ transplant (f) dilemmas, split by the set of reasons evaluated and the reported strategy used to make the initial decision.

2015; Paxton et al., 2012). Given that participants made their final decisions immediately after evaluating reasons in Study 1b, there may not have been enough time for the reasons to induce decision change in a larger proportion of participants. To test this hypothesis, Study 1c implements a one-day delay between evaluating reasons and making the final decision. Second, although partici-

pants were relatively resistant to changing their decisions in Study 1b, there might be subtler changes in decision confidence that cannot be captured through assessment with a binary choice. Decision confidence may decrease after seeing opposing reasons relative to affirming reasons. To account for this possibility, Study 1c assesses potential changes in decision confidence after evaluating different reasons. Third, if the reasons evaluated are rarely novel to participants, then it is unlikely that they would bring about decision change. In other words, if the reasons presented had already been considered by participants when they made their initial decisions, then it is unlikely that considering the same reasons again would change their minds. Thus, Study 1c accounts for the number of reasons that participants considered novel in each condition.

Table 4

Full Results From ANOVAs

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
<b>Crying baby</b>				
Condition	1	132.49	<.001	.221
Initial strategy	1	.00	.961	.000
Condition $\times$ Initial strategy	1	.24	.623	.001
<b>Concentration camp</b>				
Condition	1	25.11	<.001	.049
Initial strategy	1	1.98	.160	.004
Condition $\times$ Initial strategy	1	.52	.472	.001
<b>Organ transplant</b>				
Condition	1	249.06	<.001	.338
Initial strategy	1	1.48	.224	.003
Condition $\times$ Initial strategy	1	4.02	.045	.008

*Note.* Full results from the three separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including condition (binary factor: affirming reasons or opposing reasons) and the strategy used to make the initial decision as independent variables. The dependent variable was the average rated quality of the reasons evaluated.

## Method

**Participants.** One thousand, four hundred fifty-four individuals voluntarily completed both sessions of this study on AMT for monetary compensation; 79% of participants who completed Session 1 returned and completed Session 2 the following day. Participant recruitment was restricted to individuals in the United States with a prior approval rating above 80%. Fourteen participants failed to follow instructions, so data were analyzed with the remaining 1,440 individuals ( $M_{\text{age}} = 35.41$ ,  $SD = 11.19$ ,  $\text{range}_{\text{age}} = [18-76]$ , 672 females). This study was approved by the Duke University Campus Institutional Review Board.

**Materials and procedure.** Only the crying baby and concentration camp dilemmas were used in Study 1c (see Appendix A for details). The entire study was self-paced. In the first part of the study, participants were instructed to carefully read one moral dilemma. Participants were randomly assigned to one of the two moral dilemmas. After reading the dilemma, participants were instructed to make an initial decision regarding which option they would choose in the dilemma and to indicate how confident they were in that decision ( $-10 = \textit{not at all confident}$ ,  $10 = \textit{very confident}$ ). Then, participants were instructed to indicate whether they based their decision on (a) an immediate reaction or feeling after reading the dilemma such that no reasons affected the decision, (b) at least one reason, or (c) some other method. If a participant selected option (b), he or she indicated what reason or reasons brought about the decision. If a participant selected option (c), he or she indicated how the decision was made.

Then, participants were randomly assigned to exactly one of three between-subjects conditions: only affirming reasons were considered, only opposing reasons were considered, or all reasons for both options were considered. The order in which the reasons were presented was randomized. Participants were asked to rate how compelling each reason was for choosing that particular option in the dilemma on a scale from  $-10$  (*worst possible reason*) to  $10$  (*best possible reason*), and they indicated which reasons were novel (i.e., they identified which reasons they had never considered before seeing them in this study).

Participants returned for the second part of the study one day later, during which they were presented with the same dilemma that they had read the day before. After rereading the dilemma, participants were asked to make a final decision regarding which option they would choose. Then, participants rated how confident they were in the final decision ( $-10 = \textit{not at all confident}$ ,  $10 = \textit{very confident}$ ). Participants saw no reasons on the second day. Upon completion, participants were monetarily compensated for their time.

## Results

Table 5 shows the percentage of participants indicating they would make each possible decision in each dilemma as a function of the set of reasons presented and when the decision was made (the initial decision before seeing reasons or the final decision one

day later after having seen reasons). Furthermore, 52.25% and 48.71% of participants reported having made their initial decisions based upon at least one reason in the crying baby and concentration camp dilemmas, respectively; 46.98% and 50.53% of participants reported having made their initial decisions based upon an immediate reaction or feeling in the crying baby and concentration camp dilemmas, respectively. The remaining participants indicated that they made their decisions in some other way.

Descriptive statistics indicate that participants had typically not considered several of the reasons prior to seeing them in the study. Specifically, in the crying baby dilemma, participants indicated that, on average, 2.42 ( $SD = 1.14$ ) of 6 affirming reasons, 2.21 ( $SD = 1.18$ ) out of 6 opposing reasons, and 4.09 ( $SD = 2.45$ ) of 12 reasons for both options were novel. Similarly, in the concentration camp dilemma, participants indicated that, on average, 2.70 ( $SD = 1.23$ ) of 6 affirming reasons, 2.82 ( $SD = 1.46$ ) of 6 opposing reasons, and 4.88 ( $SD = 2.82$ ) of 12 reasons for both options were novel.

**The effect of reasons on changing moral decisions.** First, we investigated whether participants were more likely to change their initial decisions after considering affirming reasons, opposing reasons, or reasons for both options. Collapsing across both dilemmas, there was a small but statistically significant relationship between whether or not participants changed their decisions and the particular reasons evaluated (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .017$ ,  $p < .001$ ). Evaluating the dilemmas separately, there were small but statistically significant relationships between whether or not participants changed their decisions and the particular reasons evaluated in the crying baby dilemma (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .026$ ,  $p < .001$ ) and the concentration camp dilemma (Fisher's exact test:  $p = .042$ , two-sided; Goodman and Kruskal  $\tau = .010$ ,  $p = .041$ ). Figure 4 depicts the percentages of participants who changed their decisions in each dilemma as a function of the reasons considered. Relatively few participants changed their initial decisions regardless of the reasons considered.

Employing ransacking (Goodman, 1969), we found that participants who considered only opposing reasons compared with only affirming reasons were more likely to change their initial decisions in the crying baby dilemma (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .038$ ,  $p < .001$ ) and the con-

Table 5  
*Percentages of Participants Making Each Possible Decision*

Dilemma	All reasons	Affirming reasons	Opposing reasons
Initial decision			
Crying baby (smother baby)	42% ( $n = 245$ )	42% ( $n = 262$ )	48% ( $n = 272$ )
Concentration camp (pull chair)	32% ( $n = 220$ )	33% ( $n = 215$ )	36% ( $n = 226$ )
Final decision			
Crying baby (smother baby)	42% ( $n = 245$ )	42% ( $n = 262$ )	43% ( $n = 272$ )
Concentration camp (pull chair)	29% ( $n = 220$ )	33% ( $n = 215$ )	34% ( $n = 226$ )

*Note.* Percentages of participants who indicated they would choose the option specified in parentheses in each dilemma as a function of condition (affirming reasons, opposing reasons, or reasons for both options) and when the decision was made (the initial decision before seeing reasons or the final decision after seeing reasons one day later) are depicted. Specifically, we report the percentage of participants who indicated that they would choose to smother the crying baby or pull the chair out from underneath the son. The total number of participants in each condition is also reported.

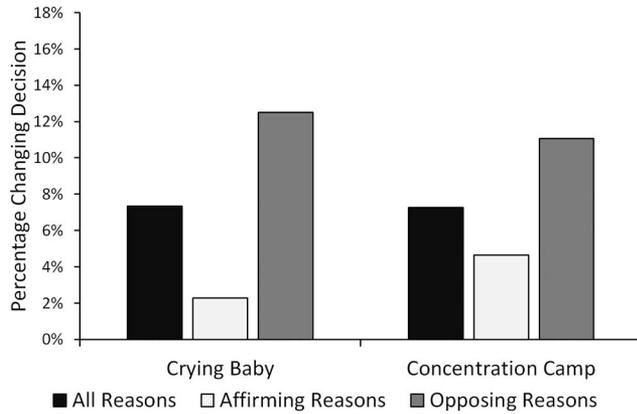


Figure 4. The percentages of participants changing their decisions in each dilemma and condition are depicted.

centration camp dilemma (Fisher's exact test:  $p = .014$ , two-sided; Goodman and Kruskal  $\tau = .014$ ,  $p = .014$ ). Participants who considered only opposing reasons compared with reasons for both options were somewhat more likely to change their initial decisions in the crying baby dilemma (Fisher's exact test:  $p = .057$ , two-sided; Goodman and Kruskal  $\tau = .007$ ,  $p = .057$ ), but there was no relationship in the concentration camp dilemma (Fisher's exact test:  $p = .191$ , two-sided; Goodman and Kruskal  $\tau = .004$ ,  $p = .191$ ). Participants who considered reasons for both options were more likely to change their initial decisions than participants who considered only affirming reasons in the crying baby dilemma (Fisher's exact test:  $p = .011$ , two-sided; Goodman and Kruskal  $\tau = .014$ ,  $p = .011$ ) but not the concentration camp dilemma (Fisher's exact test:  $p = .313$ , two-sided; Goodman and Kruskal  $\tau = .003$ ,  $p = .313$ ). All results are reported prior to correction for multiple comparisons. For both dilemmas, while participants were significantly more likely to change their decisions when only presented with opposing reasons compared with being presented with affirming reasons, all effect sizes were small. Within each condition and dilemma, no more than 14% of participants changed their initial decisions after considering reasons (Figure 4).

Next, we sought to determine the extent to which the initial decision predicts the final decision made after evaluating different reasons while also accounting for reason novelty. To this end, two separate logistic regression models were computed, one for each moral dilemma. In each model, the dependent variable is the final decision made in the moral dilemma after considering reasons, and the three independent variables are (1) the option initially chosen in the dilemma before examining reasons, (2) the reasons considered (i.e., condition: affirming reasons, opposing reasons, or reasons for both options), and (3) the total number of novel reasons for each participant. All possible interactions were also modeled. Even after accounting for the particular reasons evaluated and reason novelty, the initial decision was still a strong and significant predictor of the final decision in the crying baby dilemma ( $b = 7.85$ ,  $SE = 2.10$ ,  $e^b = 2559.46$ ,  $Z = 3.73$ ,  $p < .001$ , 95% CI = [4.12, 12.65]) and the concentration camp dilemma ( $b = 3.92$ ,  $SE = 1.78$ ,  $e^b = 50.35$ ,  $Z = 2.21$ ,  $p = .027$ , 95% CI = [.26, 7.58]). Participants typically chose the same option before and after evaluating reasons, regardless of which reasons were evaluated or

how many reasons were novel. No other variables in either model reached significance (all  $ps > .10$ ). Separate models that included only the initial decision as a predictor of the final decision produced Nagelkerke  $R^2$  values of .77 for the crying baby dilemma and .73 for the concentration camp dilemma, suggesting that the majority of variance in the final decision is explained by the initial decision.

**The effect of reasons on changes in decision confidence.** Although participants were generally resistant to changing their decisions in both dilemmas regardless of which reasons they evaluated, it is possible that the change in decision confidence from the initial decision to the final decision differs as a function of the reasons considered. To this end, we investigated whether the change in decision confidence from the initial decision to the final decision differed as a function of the reasons considered for the subset of participants who did not change their decisions after evaluating reasons. Two separate one-way ANOVAs revealed that the change in decision confidence differed as a function of condition in the crying baby dilemma ( $F(2, 718) = 6.77$ ,  $p = .001$ ,  $\eta_p^2 = .019$ ; Figure 5) and the concentration camp dilemma ( $F(2, 607) = 7.33$ ,  $p = .001$ ,  $\eta_p^2 = .024$ ; Figure 5). Surprisingly, however, regardless of the reasons evaluated, participants were more confident, on average, in their final decisions than their initial decisions (Figure 5). So, even when participants only evaluated opposing reasons, they tended to be *more* confident in their final decisions than their initial decisions, on average.

Post hoc pairwise comparisons with Bonferroni correction for each ANOVA revealed that there was a significant difference in the change in decision confidence when only affirming reasons were evaluated compared with when only opposing reasons were evaluated in the crying baby dilemma (Mean Difference = 1.28,  $SE = .36$ ,  $p = .001$ , 95% CI = [.42, 2.14]) and the concentration camp dilemma (Mean Difference = 1.70,  $SE = .48$ ,  $p = .001$ , 95% CI = [.56, 2.84]). In both dilemmas, participants became more confident in their final decisions when they only evaluated affirming reasons compared with when they only evaluated opposing reasons. Additionally, there was a significant difference in the change in decision confidence when only opposing reasons were evaluated compared with when reasons for both options were

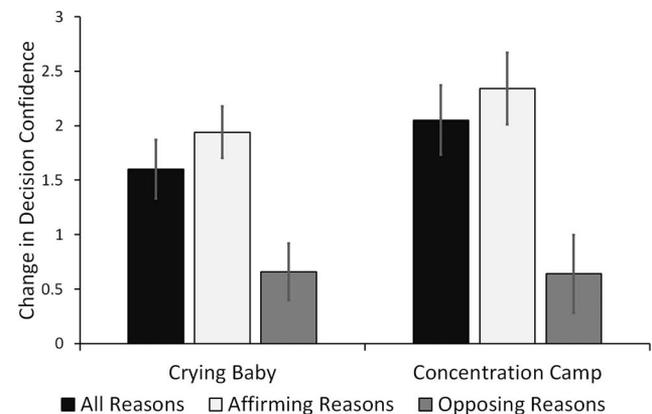


Figure 5. Means and standard error bars for the change in decision confidence (final decision confidence minus initial decision confidence) in each dilemma and condition are depicted.

evaluated in the crying baby dilemma (Mean Difference = .94,  $SE = .37$ ,  $p = .032$ , 95% CI = [.059, 1.83]) and the concentration camp dilemma (Mean Difference = 1.42,  $SE = .48$ ,  $p = .009$ , 95% CI = [.28, 2.56]). In both dilemmas, participants became more confident in their final decisions when they evaluated reasons for both options compared with when they only evaluated opposing reasons. There were no significant differences in decision confidence for either dilemma between cases where only affirming reasons were evaluated versus cases where reasons for both options were evaluated (both  $ps > .50$ ).

**Motivated evaluation of reasons.** Having found that participants rarely changed their decisions regardless of the reasons considered, we then investigated whether this resistance to changing decisions is at least partly attributable to a biased, motivated evaluation of reasons. We first investigated whether participants engaged in this kind of evaluation when considering reasons for both options in the dilemmas. Specifically, after making an initial decision and then evaluating reasons for both options, we investigated whether participants evaluated reasons supporting that decision more favorably than reasons supporting the other possible decision in the dilemma. To this end, we computed separate two-way ANOVAs for each dilemma. For each ANOVA, the initial decision (binary factor) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable for each ANOVA was computed by first averaging the rated quality of reasons for each possible decision separately and then taking the difference in average ratings for one option relative to the other option.

Our results indicate that participants do evaluate the reasons supporting their initial decisions more favorably than the reasons supporting the foregone option in each dilemma, regardless of whether the initial decision was reportedly the product of at least one reason or an immediate, affective reaction. Table 6 displays results from each two-way ANOVA. In each model, there was a large and significant main effect of initial decision. Specifically, the main effects of initial decision indicated that individuals rate reasons supporting their initial decisions more favorably than those supporting the foregone option for the crying baby dilemma (Mean difference = 8.12,  $SE = .75$ ,  $p < .001$ , 95% CI = [6.64, 9.59];

Figure 6a) and the concentration camp dilemma (Mean difference = 4.23,  $SE = .65$ ,  $p < .001$ , 95% CI = [2.94, 5.51]; Figure 6b). The significant main effect of initial decision in the crying baby dilemma was qualified by a small but significant interaction between the initial decision and initial strategy (Table 6), but subsequent tests of simple main effects showed that participants rated reasons supporting their initial decisions more favorably, regardless of the reported strategy used to make the initial decision (both  $ps < .001$ ; Figure 6).

Second, we investigated whether participants engaged in a biased, motivated reasoning when examining only affirming reasons versus only opposing reasons. Specifically, we investigated whether individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons, regardless of the reported strategy used to make the initial decision. To this end, separate two-way ANOVAs were computed for each dilemma. For each ANOVA, the condition (binary factor: only opposing reasons evaluated or only affirming reasons evaluated) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable was the average rated quality of the reasons evaluated. Table 7 displays results from each two-way ANOVA. In each model, there was a large and significant main effect of condition. Specifically, the main effects of condition indicated that individuals rated affirming reasons more favorably than opposing reasons for the crying baby dilemma (Mean difference = 3.58,  $SE = .33$ ,  $p < .001$ , 95% CI = [2.93, 4.22]; Figure 6c) and the concentration camp dilemma (Mean difference = 3.12,  $SE = .35$ ,  $p < .001$ , 95% CI = [2.44, 3.81]; Figure 6d). There were no significant interactions.

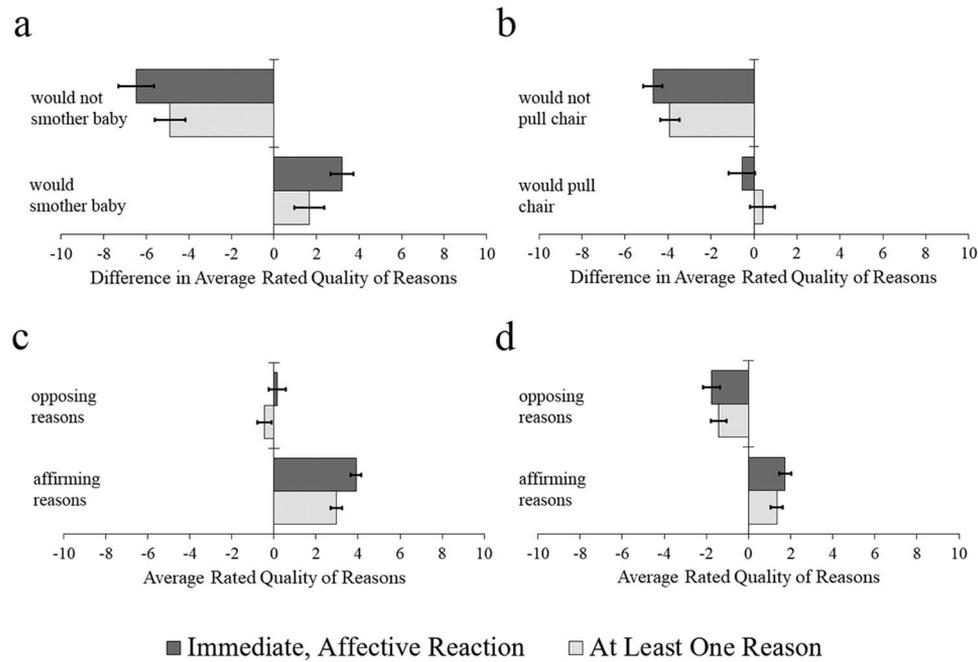
## Discussion

Consistent with the previous studies, results from Study 1c suggest that people are more likely to alter their initial decisions when presented with only opposing reasons compared with only affirming reasons. Nevertheless, the reported strength of the associations for these significant results suggest that knowing which set of reasons were considered by participants only reduces prediction error rate by less than 4%. Additional analyses revealed that participants were significantly more likely not to change their initial decisions than to change those decisions after evaluating reasons, regardless of the set of reasons considered. Again, the initial decision before seeing reasons accounted for the clear majority of the variance in predicting the final decision. These results suggest that participants' initial decisions before seeing reasons are much better predictors of the final decisions than the set of reasons considered, even after accounting for the potential effects of reason novelty and time. Participants very rarely change their initial decisions. Reason novelty and time did not affect the likelihood that participants change their decisions. Despite this unwillingness to change decisions after evaluating reasons, participants who considered only affirming reasons tended to become more confident in their final decisions relative to those who considered only opposing reasons. These effect sizes were small, however. These results replicate and extend our findings from Study 1b.

Table 6  
Full Results From ANOVAs

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
Crying baby				
Initial decision	1	117.82	<.001	.330
Initial strategy	1	.00	.969	.000
Initial Decision $\times$ Initial strategy	1	4.34	.038	.018
Concentration camp				
Initial decision	1	41.95	<.001	.163
Initial strategy	1	1.79	.182	.008
Initial Decision $\times$ Initial strategy	1	.02	.894	.000

*Note.* Full results from the two separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including the initial decision and the strategy used to make that initial decision as independent variables. The dependent variable was the difference in the average rated quality of reasons between the two possible choices.



**Figure 6.** For those participants who evaluated reasons for both options, means and standard error bars are shown for differences in the average rated quality of reasons between the two possible choices in the crying baby (a) and concentration camp (b) dilemmas, split by the initial decision and the reported strategy used to make the initial decision. Difference scores were computed by subtracting the average rating of reasons for not committing the action from the average rating of reasons for committing the action. For example, the difference in average ratings for the crying baby dilemma was computed by subtracting the average rating for not smothering the baby from the average rating for smothering the baby. For participants who evaluated only affirming reasons or only opposing reasons, means and standard error bars are depicted for the average rated quality of the reasons evaluated in the crying baby (c) and concentration camp (d) dilemmas, split by the set of reasons evaluated and the reported strategy used to make the initial decision.

Additionally, after making an initial decision and then evaluating reasons for both options, participants evaluated reasons supporting the initial decision more favorably than reasons supporting the other possible option. Similarly, individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons. These findings held regardless of whether participants reported

making their initial decisions based on at least one reason or an immediate, affective reaction. Thus, individuals appear to evaluate reasons in a motivated way to support their previous decisions. As in Study 1b, the effects regarding the motivated evaluation of reasons were large in Study 1c.

## Study 2a

Studies 1b and 1c showed that participants very rarely changed their decisions after the evaluation of the reasons. To show this, we used rather contrived sacrificial dilemmas, which have previously been used in experimental settings, designed to engender considerable conflict in the decision-making process. However, historical and anecdotal evidence from more everyday moral problems suggests that people are sometimes amenable to using reasons to arrive at moral decisions and to changing their minds about moral issues (Coles, 1986; Gilligan, 1982; Pizarro & Bloom, 2003). Thus, we wanted to see whether the previously established effects hold for more commonplace moral decisions. Studies 2 does exactly this. The purpose of Study 2a is to norm and validate a set of more everyday moral dilemmas and accompanying reasons for making each possible decision. As in Study 1a, we sought to ensure that the reasons for each option in each dilemma are relatively exhaustive.

**Table 7**  
*Full Results From ANOVAs*

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
Crying baby				
Condition	1	119.08	<.001	.185
Initial strategy	1	5.66	.018	.011
Condition $\times$ Initial strategy	1	.23	.635	.000
Concentration camp				
Condition	1	80.45	<.001	.157
Initial strategy	1	.01	.943	.000
Condition $\times$ Initial strategy	1	1.04	.308	.002

*Note.* Full results from the two separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including condition (binary factor: affirming reasons or opposing reasons) and the strategy used to make the initial decision (binary factor: as independent variables. The dependent variable was the average rated quality of the reasons evaluated.

## Method

**Participants.** Two hundred forty-six individuals voluntarily participated in this study via AMT. Participant recruitment was restricted to individuals in the United States with a prior approval rating above 85%. Sixteen participants failed to follow instructions: data were analyzed with the remaining 230 individuals ( $M_{\text{age}} = 34.44$ ,  $SD = 10.62$ ,  $\text{range}_{\text{age}} = [18-68]$ , 99 females). Racial/ethnic groups included 77.0% White non-Hispanic, 9.1% Black non-Hispanic, 6.1% Hispanic, 5.2% Asian non-Hispanic, .4% Native American, and 1.7% Other. Overall, participants in the sample were well-educated (15.6% only completed high school, 38.3% completed some college/associate degree, 37.0% completed a bachelor's degree, 6.5% completed a master's degree, and 2.6% completed a PhD). This study was approved by the Duke University Campus Institutional Review Board.

**Materials.** Three everyday moral dilemmas, each created to facilitate conflict between reasons for competing possible choices, were used. In each dilemma, there were only two possible choices that participants could make. The norming and selection process for these everyday moral dilemmas mirrored those for the philosophical dilemmas described previously. A separate group of 20 participants recruited from AMT participated in a norming session and were excluded from subsequent studies. In the norming session, the three moral scenarios were presented separately to each participant. Participants were asked to provide their best four reasons for choosing each option in each moral scenario presented herein. We compiled the seven mostly commonly reported reasons for choosing each option, which were modified for clarity and brevity. These reasons were diverse: whereas some capture distinct moral principles or obligations (e.g., "Being honest is always the best course of action"), others appealed to anticipated emotional reactions (e.g., "I might feel guilty for keeping the extra money") or monetary, self-serving reasons (e.g., "Reporting the accident might cause my insurance premium to increase"). See Appendix B for the moral scenarios and reasons used in Study 2a-c.

**Procedure.** The procedure in Study 2a closely mirrored the procedure from Study 1a. First, participants read a randomly selected moral dilemma and made judgments about seven reasons for choosing each respective option. Specifically, participants were asked to evaluate the quality of each reason for choosing that particular option in the scenario on a scale from  $-10$  (*worst possible reason*) to  $10$  (*best possible reason*). All reasons were presented on the same screen, but reasons for each option were grouped together separately. After making all 14 ratings for a given dilemma, participants were then asked to indicate which option they would choose in the scenario. Finally, participants were asked if they considered any additional reasons not presented in the study when making their decision. This same procedure was then repeated for a second randomly selected moral dilemma, and then for the third and final moral dilemma. The order of the moral dilemmas and the reasons for each option were randomized without replacement. The entire study was self-paced. Upon completion, participants were monetarily compensated for their time.

## Results and Discussion

In sum, 59.57% of participants indicated that they would choose to leave a note on the other car offering to pay for the damages in the car accident dilemma, 40.00% of participants indicated that

they would choose to use the cheat sheet on the exam in the cheating dilemma, and 70.43% of participants indicated that they would return the extra change to the clerk in the extra change dilemma. Only 8.70%, 5.65%, and 7.39% of participants reported having considered an additional reason or additional reasons beyond those provided in the car accident, cheating, and extra change dilemmas, respectively. This finding suggests that our normed set of reasons for choosing each option in the moral dilemmas is relatively exhaustive. Supplementary Figure 2 depicts distributions of ratings for each reason in each moral dilemma across participants.

## Study 2b

In Study 2b, we consider the influence of reasons on changing decisions in more everyday moral dilemmas. As in Study 1b, after making an initial choice in a moral dilemma, we present participants with a set of affirming reasons, opposing reasons, or reasons for both options. We investigate whether participants are more likely to change their decisions after considering opposing reasons versus affirming reasons. We also investigate whether the initial decision is a better predictor of the final decision than the set of reasons presented. Furthermore, we investigate whether people engage in a biased, motivated evaluation of reasons to support their initial, affective decisions.

## Method

**Participants.** Two thousand, three hundred forty-four individuals voluntarily participated in this study via AMT for monetary compensation. Participant recruitment was restricted to individuals in the United States with a prior approval rating above 80%. Forty-nine participants failed to follow instructions, so data were analyzed with the remaining 2,295 individuals ( $M_{\text{age}} = 34.94$ ,  $SD = 11.17$ ,  $\text{range}_{\text{age}} = [18-85]$ , 968 females). This study was approved by the Duke University Campus Institutional Review Board.

**Materials and procedure.** The same set of moral dilemmas and reasons for each course of action from Study 2a was used in Study 2b (see Appendix B for details). The same procedure used in Study 1b was also used in Study 2b. The only difference between Study 1b and Study 2b is the set of dilemmas and accompanying reasons.

## Results

Table 8 shows the percentage of participants indicating they would make each possible decision in each dilemma as a function of the set of reasons presented (affirming reasons, opposing reasons, or reasons for both options) and when the decision was made (the initial decision before seeing reasons or the final decision after seeing reasons). Furthermore, 40.47%, 44.47%, and 45.36% of participants reported having made their initial decisions based upon at least one reason in the car accident, cheating, and extra change dilemmas, respectively; 58.61%, 54.10%, and 53.21% of participants reported having made their initial decisions based upon an immediate reaction or feeling in the car accident, cheating, and extra change dilemmas, respectively. The remaining participants reported having made their initial decisions in some other way.

Table 8  
Percentages of Participants Making Each Possible Decision

Dilemma	All reasons	Affirming reasons	Opposing reasons
Initial decision			
Car accident (leave note)	52% ( <i>n</i> = 258)	54% ( <i>n</i> = 252)	58% ( <i>n</i> = 251)
Cheating (cheat)	50% ( <i>n</i> = 254)	42% ( <i>n</i> = 256)	45% ( <i>n</i> = 260)
Extra change (return money)	58% ( <i>n</i> = 260)	62% ( <i>n</i> = 254)	52% ( <i>n</i> = 250)
Final decision			
Car accident (leave note)	57% ( <i>n</i> = 258)	54% ( <i>n</i> = 252)	62% ( <i>n</i> = 251)
Cheating (cheat)	47% ( <i>n</i> = 254)	41% ( <i>n</i> = 256)	42% ( <i>n</i> = 260)
Extra change (return money)	64% ( <i>n</i> = 260)	63% ( <i>n</i> = 254)	57% ( <i>n</i> = 250)

*Note.* Percentages of participants who indicated they would choose the option specified in parentheses in each dilemma as a function of condition (affirming reasons, opposing reasons, or reasons for both options) and when the decision was made (the initial decision before seeing reasons or the final decision after seeing reasons) are depicted. Specifically, we report the percentage of participants who indicated that they would choose to leave a note on the other car, cheat on the exam, and return the extra money to the clerk. The total number of participants in each condition is also reported.

**The effect of reasons on changing moral decisions.** First, we sought to determine whether participants were more likely to change their initial decisions after considering affirming reasons, opposing reasons, or reasons for both options. Collapsing across all three dilemmas, there was a small but statistically significant relationship between whether or not participants changed their decisions and the particular reasons evaluated (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .017$ ). Evaluating the dilemmas separately, there was a statistically significant relationship between whether or not participants changed their decisions and the particular reasons evaluated in the car accident (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .024$ ,  $p < .001$ ), cheating (Fisher's exact test:  $p = .049$ , two-sided; Goodman and Kruskal  $\tau = .007$ ,  $p = .084$ ), and extra change (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .023$ ,  $p < .001$ ) dilemmas. Figure 7 depicts the percentages of participants who changed their decisions in each dilemma as a function of the particular reasons considered.

Employing ransacking (Goodman, 1969), we found that participants who considered only opposing reasons were more likely to change their initial decisions than participants who considered only affirming reasons in the car accident (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .038$ ,  $p < .001$ ),

cheating (Fisher's exact test:  $p = .033$ , two-sided; Goodman and Kruskal  $\tau = .011$ ,  $p = .033$ ), and extra change (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .035$ ,  $p < .001$ ) dilemmas. Participants who considered only opposing reasons were not significantly more likely to change their initial decisions compared with participants who considered reasons for both options in the car accident dilemma (Fisher's exact test:  $p = .395$ , two-sided; Goodman and Kruskal  $\tau = .002$ ,  $p = .395$ ) or the cheating dilemma (Fisher's exact test:  $p = .828$ , two-sided; Goodman and Kruskal  $\tau = .000$ ,  $p = .828$ ), but there was a trend in the extra change dilemma (Fisher's exact test:  $p = .085$ , two-sided; Goodman and Kruskal  $\tau = .007$ ,  $p = .085$ ). Participants who considered reasons for both options were significantly more likely to change their initial decisions than participants who considered only affirming reasons in the car accident dilemma (Fisher's exact test:  $p < .001$ , two-sided; Goodman and Kruskal  $\tau = .026$ ,  $p < .001$ ) and the extra change dilemma (Fisher's exact test:  $p = .015$ , two-sided; Goodman and Kruskal  $\tau = .013$ ,  $p = .015$ ), but there was only a marginally significant relationship in the cheating dilemma (Fisher's exact test:  $p = .054$ , two-sided; Goodman and Kruskal  $\tau = .008$ ,  $p = .088$ ). All results are reported prior to correction for multiple comparisons. Across all three dilemmas, while participants were significantly more likely to change their decisions when only presented with opposing reasons compared with being presented with affirming reasons, all effect sizes were small. Within each condition and dilemma, no more than 12% of participants changed their initial decisions (Figure 7).

Next, we sought to determine the extent to which the initial decision predicts the final decision. To this end, three separate logistic regression models were computed, one for each moral dilemma. In each model, the dependent variable is the final decision made in the moral dilemma after considering reasons, and the two independent variables are (1) the option originally chosen in the dilemma before examining reasons and (2) the reasons considered (i.e., condition: affirming reasons, opposing reasons, or reasons for both options). Even after accounting for the particular reasons evaluated, the initial decision was a strong and significant predictor of the final decision in the car accident ( $b = 5.37$ ,  $SE = .32$ ,  $e^b = 215.57$ ,  $Z = 16.96$ ,  $p < .001$ , 95% CI = [4.78, 6.03]), cheating ( $b = 7.37$ ,  $SE = .53$ ,  $e^b = 1584.00$ ,  $Z = 13.81$ ,  $p < .001$ ,

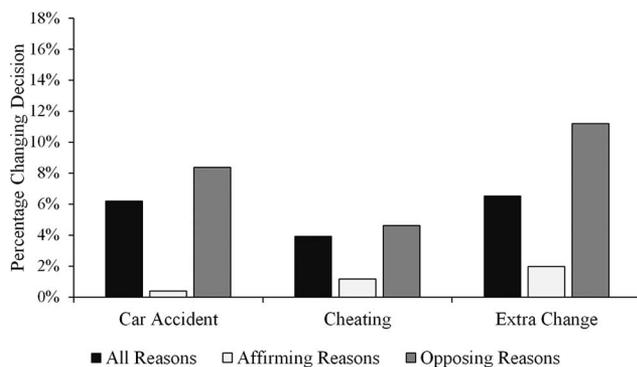


Figure 7. The percentages of participants changing their decisions in each dilemma and condition are depicted.

95% CI = [6.42, 8.55]), and extra change ( $b = 5.91$ ,  $SE = .39$ ,  $e^b = 366.90$ ,  $Z = 15.21$ ,  $p < .001$ , 95% CI = [5.20, 6.73]) dilemmas. In separate follow-up models isolating the subset of participants who considered only opposing reasons, the initial decision remained strong predictor of the final decision in all dilemmas (all  $ps < .001$ ); participants rarely changed their decisions from the initial choice to the final choice even in the face of opposing reasons (Figure 7). Additionally, regardless of which reasons were considered, participants reported high confidence in their final decisions (supplementary Table 2).

We next sought to compare pseudo- $R^2$  values between models that included versus did not include condition (i.e., affirming reasons, opposing reasons, or reasons for both options) to determine how much variance each condition explains above and beyond the initial decision. After running three additional logistic regression models excluding condition as an independent variable, the amount of variance explained in the models for each dilemma changed minimally compared with the full models with all predictors. The models that only included the initial decision accounted for most of the variance in the final decision made. Table 9 reports these pseudo- $R^2$  values for each model.

**Motivated evaluation of reasons.** Having found that participants rarely changed their decisions regardless of which reasons were considered, we then investigated whether this resistance to changing decisions is at least partly attributable to a biased, motivated evaluation of reasons. To do so, we first investigated whether participants engaged in a biased, motivated evaluation of reasons when evaluating all reasons for both possible options in the dilemmas. Specifically, after making an initial decision and then evaluating reasons for both options, we investigated whether participants evaluated reasons supporting that decision more favorably than reasons supporting the other possible decision in the dilemma. To this end, separate two-way ANOVAs were computed for each dilemma. For each ANOVA, the initial decision (binary factor) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable for each ANOVA was computed by first averaging the rated quality of reasons for each possible decision separately and then taking the difference in average ratings for one option relative to the other option.

Our results indicate that participants evaluate the reasons supporting their initial decisions more favorably than the reasons supporting the foregone option in each moral dilemma, regardless of whether the initial decision was reportedly the product of at least one reason or an immediate, affective reaction. Table 10

Table 9  
Model Comparisons

Model	Car accident	Cheating	Extra change
Model with condition	.784	.896	.805
Model without condition	.781	.895	.801

*Note.* Nagelkerke  $R^2$  values are depicted for each of six different logistic regression models. For each dilemma, one logistic regression model was computed including condition (i.e., a categorical variable with levels for affirming reasons, opposing reasons, or reasons for both options) and the initial decision as predictors of the final decision, while another logistic regression model was run only including the initial decision as a predictor of the final decision.

Table 10  
Full Results From ANOVAs

Dilemma	df	F	p	Partial $\eta^2$
Car accident				
Initial decision	1	139.32	<.001	.356
Initial strategy	1	.02	.889	.000
Initial Decision $\times$ Initial strategy	1	.82	.367	.003
Cheating				
Initial decision	1	157.07	<.001	.389
Initial strategy	1	.02	.889	.000
Initial Decision $\times$ Initial strategy	1	1.36	.245	.005
Extra change				
Initial decision	1	165.27	<.001	.396
Initial strategy	1	.49	.485	.002
Initial Decision $\times$ Initial strategy	1	1.55	.215	.006

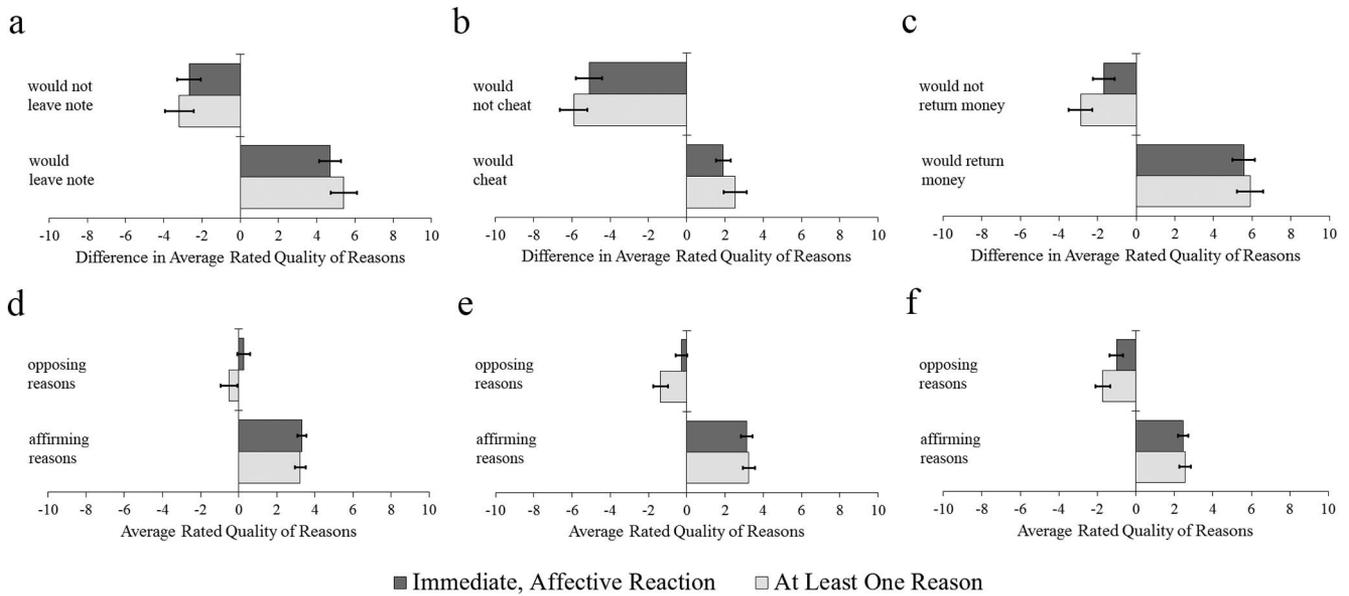
*Note.* Full results from the three separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including the initial decision and the strategy used to make that initial decision as independent variables. The dependent variable was the difference in the average rated quality of reasons between the two possible choices.

displays results from each two-way ANOVA. In each model, we found a significant main effect of initial decision, but no effect of initial strategy and no interaction between initial decision and initial strategy. Specifically, the main effects of initial decision indicated that individuals rate reasons supporting their initial decisions more favorably than those supporting the foregone option for the car accident dilemma (Mean difference = 7.99,  $SE = .68$ ,  $p < .001$ , 95% CI = [6.66, 9.32]; Figure 8a), the cheating dilemma (Mean difference = 7.72,  $SE = .62$ ,  $p < .001$ , 95% CI = [6.51, 8.94]; Figure 8b), and the extra change dilemma (Mean difference = 8.01,  $SE = .62$ ,  $p < .001$ , 95% CI = [6.78, 9.24]; Figure 8c). There were no significant interactions in any of the models.

Second, we investigated whether participants engaged in a biased, motivated evaluation of reasons when evaluating only affirming reasons versus only opposing reasons. Specifically, we investigated whether individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons. To this end, separate two-way ANOVAs were computed for each dilemma. For each ANOVA, the condition (binary factor: only opposing reasons evaluated or only affirming reasons evaluated) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable was the average rated quality of reasons evaluated. Table 11 displays results from each two-way ANOVA. In each model, there was a significant main effect of condition. Specifically, the main effects of condition indicated that individuals rated affirming reasons more favorably than opposing reasons in the car accident (Mean difference = 3.40,  $SE = .33$ ,  $p < .001$ , 95% CI = [2.76, 4.05]; Figure 8d), cheating (Mean difference = 4.02,  $SE = .33$ ,  $p < .001$ , 95% CI = [3.38, 4.66]; Figure 8e), and extra change (Mean difference = 3.86,  $SE = .33$ ,  $p < .001$ , 95% CI = [3.21, 4.51]; Figure 8f) dilemmas.

## Discussion

Across all three moral dilemmas, our results suggest that people are more likely to change their initial decisions when



*Figure 8.* For those participants who evaluated reasons for both options, means and standard error bars are depicted for differences in the average rated quality of reasons between the two possible choices in the car accident (a), cheating (b), and extra change (c) dilemmas, split by the initial decision and the reported strategy used to make the initial decision. Difference scores were computed by subtracting the average rating of reasons for not committing the action from the average rating of reasons for committing the action. For example, the difference in average ratings for the cheating dilemma was computed by subtracting the average rating for not cheating on the exam from the average rating for cheating on the exam. For participants who evaluated only affirming reasons or only opposing reasons, means and standard error bars are depicted for the average rated quality of the reasons evaluated in the car accident (d), cheating (e), and extra change (f) dilemmas, split by the set of reasons evaluated and the reported strategy used to make the initial decision.

presented with only opposing reasons compared with only affirming reasons. Nevertheless, the reported strength of the associations for these significant results suggest that knowing which set of reasons were considered by participants only reduces prediction error rate by less than 4%. Additional analyses revealed that participants were significantly more likely

not to change their initial decisions than to change those decisions after evaluating reasons, regardless of whether participants considered affirming reasons, opposing reasons, or reasons for both options. The initial decision before seeing reasons accounted for the clear majority of the variance in predicting the final decision; the particular set of reasons considered only minimally increased goodness of fit (i.e., pseudo- $R^2$  values) when added to the model. These results suggest that participants' initial decisions before seeing reasons are much better predictors of the final decisions than the particular reasons considered. Participants very rarely change their initial decisions. These results from Study 2b closely mirror those from Study 1b.

Additionally, across all three dilemmas, after making an initial decision and then evaluating reasons for both options, participants evaluated reasons supporting the initial decision more favorably than reasons supporting the other possible option in the dilemma. Similarly, individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons. These findings held regardless of whether participants reported making their initial decisions based on at least one reason or an immediate, affective reaction. Thus, individuals do seem to evaluate reasons in a way that supports their previous decisions. This behavior is likely one important factor explaining why individuals rarely change their moral decisions in the face of reasons.

Table 11  
Full Results From ANOVAs

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
<b>Car accident</b>				
Condition	1	108.10	<.001	.180
Initial strategy	1	1.72	.190	.003
Condition $\times$ Initial strategy	1	1.12	.290	.002
<b>Cheating</b>				
Condition	1	151.30	<.001	.231
Initial strategy	1	2.20	.139	.004
Condition $\times$ Initial strategy	1	3.44	.064	.007
<b>Extra change</b>				
Condition	1	136.21	<.001	.216
Initial strategy	1	.79	.375	.002
Condition $\times$ Initial strategy	1	1.51	.220	.003

*Note.* Full results from the three separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including condition (binary factor: affirming reasons or opposing reasons) and the strategy used to make the initial decision as independent variables. The dependent variable was the average rated quality of the reasons evaluated.

### Study 2c

Building upon the results from Study 2b, we further consider the influence of reasons on changing moral decisions in Study 2c. As in Study 2b, we use everyday moral dilemmas to investigate whether people are more likely to change their initial decisions after considering opposing reasons or affirming reasons, whether the initial decision is a better predictor of the final decision than the particular set of reasons presented, and whether people engage in a biased, motivated evaluation of reasons to support their initial, affective reactions. The basic design and overarching questions in Study 2c are the same as those in Study 1c.

The same three additional variables investigated in Study 1c are also investigated in Study 2c using the everyday dilemmas. That is, (a) we implement a one day delay between evaluating reasons and making the final decision, (b) we investigate potential changes in decision confidence as a function of the reasons considered, and (c) we account for the potential influence of reason novelty in decision change.

### Method

**Participants.** One thousand, four hundred fifty-two individuals voluntarily completed both sessions of this study on AMT for monetary compensation; 82% of participants who completed Session 1 returned and completed Session 2 the following day. Participant recruitment was restricted to individuals in the United States with a prior approval rating above 80%. Seven participants failed to follow instructions, so data were analyzed with the remaining 1,445 individuals ( $M_{\text{age}} = 36.19$ ,  $SD = 11.34$ ,  $\text{range}_{\text{age}} = [18-74]$ , 740 females). This study was approved by the Duke University Campus Institutional Review Board.

**Materials and procedure.** Only the car accident and cheating dilemmas were used in Study 2c (see Appendix B for details). The procedure in Study 2c was identical to the procedure in Study 1c. The differences between Study 1c and Study 2c are the set of dilemmas and the accompanying reasons.

### Results

Table 12 shows the percentage of participants indicating they would make each possible decision in each dilemma as a function of the set of reasons presented and when the decision was made

(initial decision before seeing reasons or final decision one day later after having seen reasons). Furthermore, 40.00% and 48.30% of participants reported having made their initial decisions based upon at least one reason in the car accident and cheating dilemmas, respectively; 58.87% and 51.02% of participants reported having made their initial decisions based upon an immediate reaction or feeling in the car accident and cheating dilemmas, respectively. The remaining participants indicated that they made their decisions in some other way.

Descriptive statistics indicate that participants had typically not considered several of the reasons prior to seeing them in the study. Specifically, in the car accident dilemma, participants indicated that, on average, 2.72 ( $SD = 1.46$ ) of 7 affirming reasons, 2.77 ( $SD = 1.60$ ) of 7 opposing reasons, and 4.26 ( $SD = 2.98$ ) of 14 reasons for both options were novel. Similarly, in the cheating dilemma, participants indicated that, on average, 2.88 ( $SD = 1.53$ ) of 7 affirming reasons, 2.58 ( $SD = 1.59$ ) of 7 opposing reasons, and 4.39 ( $SD = 3.11$ ) of 14 reasons for both options were novel.

**The effect of reasons on changing moral decisions.** First, we investigated whether participants were more likely to change their initial decisions after considering affirming reasons, opposing reasons, or reasons for both options. Collapsing across both dilemmas, there was a small but statistically significant relationship between whether or not participants changed their decisions and the particular reasons evaluated (Fisher's exact test:  $p = .001$ , two-sided; Goodman and Kruskal  $\tau = .011$ ,  $p = .001$ ). Evaluating the dilemmas separately, there was a small but statistically significant relationship between whether or not participants changed their decisions and the particular reasons evaluated in the car accident dilemma (Fisher's exact test:  $p = .007$ , two-sided; Goodman and Kruskal  $\tau = .015$ ,  $p = .005$ ) and a marginally significant relationship in the cheating dilemma (Fisher's exact test:  $p = .068$ , two-sided; Goodman and Kruskal  $\tau = .008$ ,  $p = .069$ ). Figure 9 depicts the percentages of participants who changed their decisions in each dilemma as a function of the reasons considered. Relatively few participants changed their initial decisions regardless of which reasons were considered.

Employing ransacking (Goodman, 1969), we found that participants who considered only opposing reasons compared with only affirming reasons were significantly more likely to change their initial decisions in the car accident dilemma (Fisher's exact test:  $p = .006$ , two-sided; Goodman and Kruskal  $\tau = .017$ ,  $p = .006$ )

Table 12  
*Percentages of Participants Making Each Possible Decision*

Dilemma	All reasons	Affirming reasons	Opposing reasons
Initial decision			
Car accident (leave note)	60% ( $n = 242$ )	58% ( $n = 238$ )	55% ( $n = 230$ )
Cheating (cheat)	38% ( $n = 231$ )	44% ( $n = 248$ )	36% ( $n = 256$ )
Final decision			
Car accident (leave note)	60% ( $n = 242$ )	60% ( $n = 238$ )	57% ( $n = 230$ )
Cheating (cheat)	38% ( $n = 231$ )	44% ( $n = 248$ )	34% ( $n = 256$ )

*Note.* Percentages of participants who indicated they would choose the option specified in parentheses in each dilemma as a function of condition (affirming reasons, opposing reasons, or reasons for both options) and when the decision was made (the initial decision before seeing reasons or the final decision after seeing reasons) are depicted. Specifically, we report the percentage of participants who indicated that they would choose to leave a note on the other car and cheat on the exam. The total number of participants in each condition is also reported.

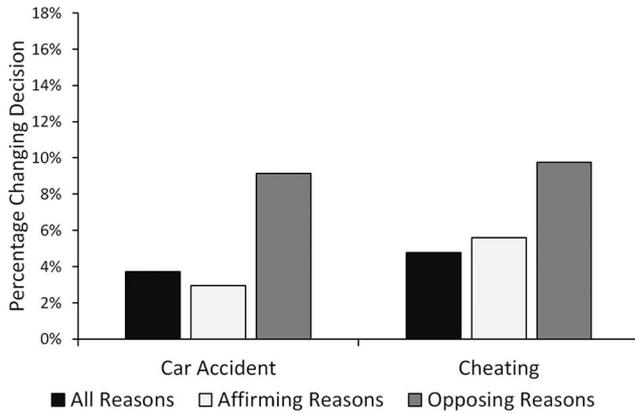


Figure 9. The percentages of participants changing their decisions in each dilemma and condition are depicted.

and somewhat more likely in the cheating dilemma (Fisher's exact test:  $p = .096$ , two-sided; Goodman and Kruskal  $\tau = .006$ ,  $p = .096$ ). Participants who considered only opposing reasons compared with reasons for both options were significantly more likely to change their initial decisions in the car accident dilemma (Fisher's exact test:  $p = .022$ , two-sided; Goodman and Kruskal  $\tau = .012$ ,  $p = .022$ ) and the cheating dilemma (Fisher's exact test:  $p = .038$ , two-sided; Goodman and Kruskal  $\tau = .009$ ,  $p = .038$ ). Participants who considered reasons for both options were not more or less likely to change their initial decisions than participants who considered only affirming reasons in the car accident dilemma (Fisher's exact test:  $p = .800$ , two-sided; Goodman and Kruskal  $\tau = .000$ ,  $p = .800$ ) or the cheating dilemma (Fisher's exact test:  $p = .687$ , two-sided; Goodman and Kruskal  $\tau = .000$ ,  $p = .999$ ). All results are reported prior to correction for multiple comparisons. For both dilemmas, while participants were at least somewhat more likely to change their decisions when only presented with opposing reasons compared with being presented with affirming reasons, all effect sizes were small. Within each condition and dilemma, no more than 10% of participants changed their initial decisions after considering reasons (Figure 9).

Next, we sought to determine the extent to which the initial decision predicts the final decision made after evaluating different reasons while also accounting for reason novelty. To this end, two separate logistic regression models were computed, one for each moral dilemma. In each model, the dependent variable is the final decision made in the moral dilemma after considering reasons, and the three independent variables are (1) the option initially chosen in the dilemma before examining reasons, (2) the reasons considered (i.e., condition: affirming reasons, opposing reasons, or reasons for both options), and (3) the total number of novel reasons for each participant. All possible interactions were also modeled. Even after accounting for the particular reasons evaluated and reason novelty, the initial decision was still a strong and significant predictor of the final decision in the car accident dilemma ( $b = 7.59$ ,  $SE = 2.26$ ,  $e^b = 1986.52$ ,  $Z = 3.36$ ,  $p < .001$ , 95% CI = [3.76, 12.76]) and the cheating dilemma ( $b = 5.74$ ,  $SE = 1.20$ ,  $e^b = 310.49$ ,  $Z = 4.79$ ,  $p < .001$ , 95% CI = [3.52, 8.25]). Participants typically chose the same option before and after evaluating reasons, regardless of which reasons were evaluated or

how many of the reasons were novel. No other variables in either model reached significance (all  $ps > .10$ ). Separate models that included only the initial decision as a predictor of the final decision produced Nagelkerke  $R^2$  values of .73 for the car accident dilemma and .77 for the cheating dilemma, suggesting that most of the variance in the final decision is explained by the initial decision.

**The effect of reasons on changes in decision confidence.** Although participants were generally resistant to changing their decisions in both dilemmas regardless of which reasons they evaluated, it is possible that the change in decision confidence from the initial decision to the final decision differs as a function of the reasons considered. To this end, we investigated whether the change in decision confidence from the initial decision to the final decision differed as a function of the reasons considered for the subset of participants who did not change their decisions after evaluating reasons. Two separate one-way ANOVAs revealed that the change in decision confidence differed significantly as a function of condition in the car accident dilemma ( $F(2, 670) = 13.75$ ,  $p < .001$ ,  $\eta_p^2 = .039$ ; Figure 10) but not in the cheating dilemma ( $F(2, 682) = 1.09$ ,  $p = .336$ ,  $\eta_p^2 = .003$ ; Figure 10). Surprisingly, however, regardless of the reasons evaluated, participants were more confident, on average, in their final decisions than in their initial decisions for the car accident dilemma (Figure 10). So, even when participants only evaluated opposing reasons, they tended to be *more* confident in their final decisions than their initial decisions, on average.

Post hoc pairwise comparisons with Bonferroni correction revealed that there was a significant difference in the change in decision confidence when only affirming reasons were evaluated compared with when only opposing reasons were evaluated in the car accident dilemma (Mean Difference = 1.75,  $SE = .33$ ,  $p < .001$ , 95% CI = [.95, 2.55]). Participants became more confident in their final decisions when they only evaluated affirming reasons compared with when they only evaluated opposing reasons. Additionally, there was a significant difference in the change in decision confidence when only opposing reasons were evaluated compared with when reasons for both options were evaluated in the car accident dilemma (Mean Difference = 1.02,  $SE = .33$ ,  $p = .007$ , 95% CI = [.22, 1.82]). Participants became more confident

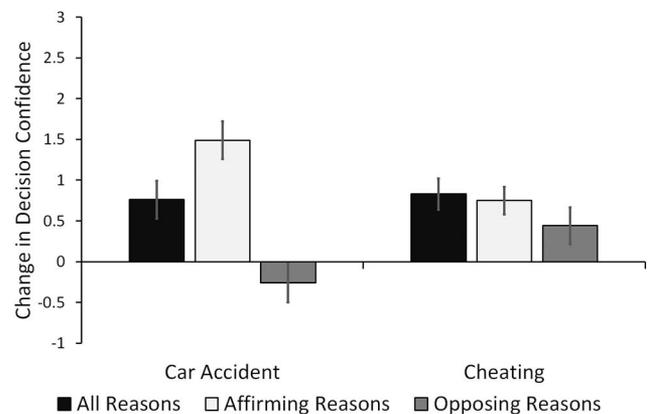


Figure 10. Means and standard error bars for the change in decision confidence (final decision confidence minus initial decision confidence) in each dilemma and condition are depicted.

in their final decisions when they evaluated reasons for both options compared with when they only evaluated opposing reasons. Finally, there was somewhat of a difference in the change in decision confidence when reasons for both options were evaluated compared with when only affirming reasons were evaluated in the car accident dilemma (Mean Difference = .73,  $SE = .33$ ,  $p = .075$ , 95% CI = [-.05, 1.51]). Participants became somewhat more confident in their final decisions when they evaluated only affirming reasons compared with when they evaluated reasons for both options.

**Motivated evaluation of reasons.** Having found that participants rarely changed their decisions regardless of the reasons considered, we then investigated whether this resistance to changing decisions is at least partly attributable to a biased, motivated evaluation of reasons. To do so, we first investigated whether participants engaged in this kind of evaluation when considering reasons for both options in the dilemmas. Specifically, after making an initial decision and then evaluating reasons for both options, we investigated whether participants evaluated reasons supporting that decision more favorably than reasons supporting the other possible decision in the dilemma. To this end, we computed separate two-way ANOVAs for each dilemma. For each ANOVA, the initial decision (binary factor) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable for each ANOVA was computed by first averaging the rated quality of reasons for each possible decision separately and then taking the difference in average ratings for one option relative to the other option.

Our results indicate that participants do evaluate the reasons supporting their initial decisions more favorably than the reasons supporting the foregone option in each dilemma, regardless of whether the initial decision was reportedly the product of at least one reason or an immediate, affective reaction. Table 13 displays results from each two-way ANOVA. In each model, there was a large and significant main effect of initial decision, but no effect of initial strategy and no interaction between initial decision and initial strategy. Specifically, the large main effects of initial decision indicated that individuals rate reasons supporting their initial decisions more favorably than those supporting the foregone option for the car accident dilemma (Mean difference = 7.43,  $SE =$

.75,  $p < .001$ , 95% CI = [5.95, 8.90]; Figure 11a) and the cheating dilemma (Mean difference = 8.17,  $SE = .70$ ,  $p < .001$ , 95% CI = [6.80, 9.55]; Figure 11b). There were no significant interactions.

Second, we investigated whether participants engaged in a biased, motivated reasoning when examining only affirming reasons versus only opposing reasons. Specifically, we investigated whether individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons, regardless of the reported strategy used to make the initial decision. To this end, separate two-way ANOVAs were computed for each dilemma. For each ANOVA, the condition (binary factor: only opposing reasons evaluated or only affirming reasons evaluated) and the reported strategy used to make the initial decision (binary factor: at least one reason vs. an immediate, affective reaction) were modeled as independent variables. The dependent variable was the average rated quality of the reasons evaluated. Table 14 displays results from each two-way ANOVA. In each model, there was a large and significant main effect of condition. Specifically, the main effects of condition indicated that individuals rated affirming reasons more favorably than opposing reasons for the car accident dilemma (Mean difference = 4.08,  $SE = .33$ ,  $p < .001$ , 95% CI = [3.44, 4.72]; Figure 11c) and the cheating dilemma (Mean difference = 4.40,  $SE = .34$ ,  $p < .001$ , 95% CI = [3.72, 5.07]; Figure 11d). There were no significant interactions.

## Discussion

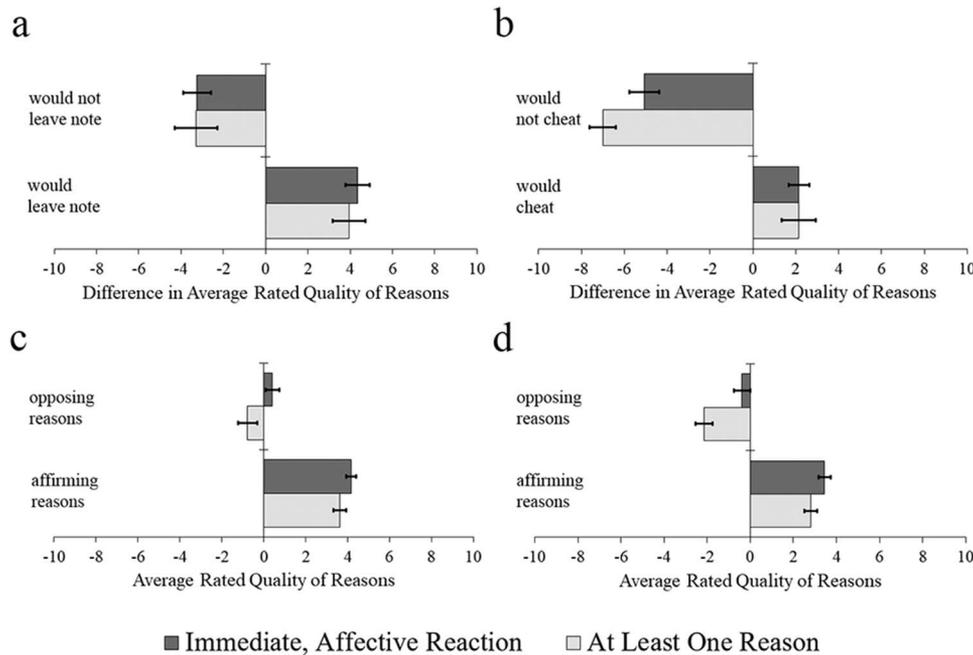
Consistent with results from the previous studies, results from Study 2c suggest that people are more likely to alter their initial decisions when presented with only opposing reasons compared with only affirming reasons. Nevertheless, the reported strength of the associations for these significant results suggest that knowing which set of reasons were considered by participants only reduces prediction error rate by less than 2%. Additional analyses revealed that participants were significantly more likely not to change their initial decisions than to change those decisions after evaluating reasons, regardless of which reasons were considered. The initial decision made before seeing reasons accounted for the clear majority of the variance in predicting the final decision. These results suggest that participants' initial decisions before seeing reasons are much better predictors of the final decisions than the set of reasons considered, even after accounting for the potential effects of reason novelty and time. Reason novelty and time did not affect the likelihood that participants change their decisions. Participants very rarely change their initial decisions. Despite this unwillingness to change decisions after evaluating reasons, participants who considered only affirming reasons tended to become more confident in their final decisions relative to those who considered only opposing reasons in the car accident dilemma but not the cheating dilemma. The significant effect in the car accident dilemma was relatively small, however.

Additionally, after making an initial decision and then evaluating reasons for both options, participants evaluated reasons supporting the initial decision more favorably than reasons supporting the other possible option. Similarly, individuals who evaluated only opposing reasons rated those reasons less favorably, on average, than individuals who evaluated only affirming reasons. These findings held regardless of whether participants reported

Table 13  
Full Results From ANOVAs

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
Car accident				
Initial decision	1	98.85	<.001	.298
Initial strategy	1	.09	.771	.000
Initial Decision $\times$ Initial strategy	1	.06	.813	.000
Cheating				
Initial decision	1	137.20	<.001	.379
Initial strategy	1	1.95	.164	.009
Initial Decision $\times$ Initial strategy	1	1.88	.172	.008

*Note.* Full results from the two separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including the initial decision and the strategy used to make that initial decision as independent variables. The dependent variable was the difference in the average rated quality of reasons between the two possible choices.



*Figure 11.* For those participants who evaluated reasons for both options, means and standard error bars are depicted for differences in the average rated quality of reasons between the two possible choices in the car accident (a) and cheating (b) dilemmas, split by the initial decision and the reported strategy used to make the initial decision. Difference scores were computed by subtracting the average rating of reasons for not committing the action from the average rating of reasons for committing the action. For example, the difference in average ratings for the cheating dilemma was computed by subtracting the average rating for not cheating on the exam from the average rating for cheating on the exam. For participants who evaluated only affirming reasons or only opposing reasons, means and standard error bars are depicted for the average rated quality of the reasons evaluated in the car accident (c) and cheating (d) dilemmas, split by the set of reasons evaluated and the reported strategy used to make the initial decision.

making their initial decisions based on at least one reason or an immediate, affective reaction. And the effect sizes were all relatively large. Thus, individuals do seem to evaluate reasons in a motivated way to support their previous decisions. As in our previous studies, the effects regarding the motivated evaluation of reasons were large in Study 1c.

**Table 14**  
*Full Results From ANOVAs*

Dilemma	<i>df</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
Car accident				
Condition	1	156.73	<.001	.254
Initial strategy	1	6.98	.009	.015
Condition $\times$ Initial strategy	1	.96	.328	.002
Cheating				
Condition	1	164.50	<.001	.249
Initial strategy	1	12.24	.001	.024
Condition $\times$ Initial strategy	1	2.72	.100	.005

*Note.* Full results from the two separate ANOVAs are depicted. For each dilemma, a two-way ANOVA was computed including condition (binary factor: affirming reasons or opposing reasons) and the strategy used to make the initial decision as independent variables. The dependent variable was the average rated quality of the reasons evaluated.

### General Discussion

Many philosophers have argued that making and revising moral decisions ought to be a matter of deliberating over relevant reasons (e.g., Broome, 2013; Crisp, 2005; Parfit, 2011; Philips, 1987; Ross, 1930; Scanlon, 1998). We investigated three descriptive questions about the role of reasons in changing decisions in diverse moral dilemmas. First, to what extent does examining affirming reasons, opposing reasons, or reasons for both options lead people to alter their initial decisions? For all dilemmas, we found that people were more likely to alter their initial decisions when presented with only opposing reasons compared with only affirming reasons. Nevertheless, the strength of the associations for the significant results suggests that knowing which set of reasons were considered by participants at best reduces prediction error rate by less than 6%. Relatively few people changed their decisions, regardless of which reasons were considered. Second, does the initial decision better predict the final decision compared with the particular reasons considered? For all dilemmas, we found that the initial decision was a much better predictor of the final decision than the set of reasons presented. In other words, most participants made the same decision after considering reasons, regardless of which reasons were considered. This finding held even after accounting for reason novelty and increasing the delay between the

initial decision and the final decision. Third, do participants rate the reasons supporting their initial decisions more favorably than reasons opposing their initial decisions, regardless of whether they report having made their initial decisions based on an immediate, affective reaction or at least one reason? For all dilemmas, no matter how participants reported making their initial decisions, we found that the reasons supporting the initial decision were judged more favorably than reasons supporting the other option.

From research in behavioral economics and consumer choice, there is evidence that attitudes, preferences, and decisions are often made or modified by the consideration and weighing of reasons (Barber et al., 2003; Petty & Cacioppo, 1986; Shafir et al., 1993; Simonson, 1989; Vlaev et al., 2009; Wilson et al., 1989). Given these findings, it may seem surprising that so few people changed their decisions when presented with opposing reasons in moral dilemmas. Perhaps there is something unique about *moral decision making*. When people view issues in terms of what is morally right and wrong, they tend to be more resistant to compromising (Skitka, 2010; Skitka, Bauman, & Sargis, 2005; Turiel, 2002). Despite the considerable variability in moral judgments and beliefs across persons, people tend to believe that others would universally agree with their moral beliefs if only they were aware of “the facts” (Morgan, Skitka, & Wisneski, 2010; Skitka, 2010; Skitka et al., 2005). And morality plays a central role in the construction and perception of the personal identity (Stanley, Henne, Iyengar, Sinnott-Armstrong, & De Brigard, 2017; Strohminger & Nichols, 2014). Given this work, it is likely that participants came into this study with strong, fixed views about what is right and wrong. Moreover, these views were likely central to participants’ identities and seen as absolute and true. Such views need not be held consciously or explicitly; they might be held implicitly, difficult to articulate with language, or manifested through immediate reactions to certain scenarios. If this is the case, then it is unsurprising that so many participants were resistant to using the reasons provided to change their decisions in these moral dilemmas.

There is another compatible explanation for our results from research on motivated reasoning and post hoc rationalization. Research on everyday reasoning suggests that people often make an initial decision and then set out to confirm the validity of that decision through biased and motivated reasoning (Kuhn, 1991; Kunda, 1990; Pyszczynski & Greenberg, 1987). For example, people do tend to apply less rigorous judgmental standards to preference-consistent than preference-inconsistent information in making medical decisions (Ditto & Lopez, 1992; Ditto et al., 1998, 2003), and people assign the greatest weight to factors that are most consistent with preference-consistent conclusions when deciding which candidate to hire for a job (Norton et al., 2004; Uhlmann & Cohen, 2005).

Some have argued that biased and motivated post hoc reasoning to further support or confirm previously held beliefs is the most prevalent kind of reasoning in moral judgment and decision making (Haidt, 2001, 2012). Similarly, the biased, motivated evaluation of the diverse reasons presented in our study at least partly explains some of our findings. Participants may have a propensity to readily dismiss those reasons for the option not initially chosen and readily agree with those reasons for the option initially chosen. Accordingly, they might justify their final choices by claiming that the reasons supporting their initial decisions were just better than the reasons for the fore-

gone option. Evaluating reasons with a preference for their previous decisions would make it unlikely that participants change their minds. Our results do provide direct empirical support for this explanation. After making an initial decision—regardless of whether participants reported making the initial decision based on an immediate, affective reaction or at least one reason—participants rated opposing reasons less favorably than affirming reasons. Participants then rarely changed their decisions after evaluating reasons, regardless of how they reported making their initial decisions.

Philosophers have been primarily concerned with the prescriptive projects of determining what it means to say that we have the most reason to do an act (Evers, 2010), how reasons should be weighed in so far as they count as evidence (Kearns & Star, 2009, 2013; Philips, 1987), and how weights ascribed to different reasons should accrue when making a moral decision (Nair, 2016). In contrast, our project is descriptive. We investigate the extent to which people do, in practice, use available reasons to revise their decisions in moral dilemmas. If these prescriptive ideas provide insights about how one ought to make and revise moral decisions, people should be open to changing their moral judgments and decisions as new reasons come to bear on their circumstances. Future empirical research will investigate (a) the specific conditions under which people are most likely to change their moral decisions after considering relevant reasons and (b) how to best encourage people to evaluate reasons in a less biased fashion.

## References

- Barber, B. M., Heath, C., & Odean, T. (2003). Good reasons sell: Reason-based choice among group and individual investors in the stock market. *Management Science*, *49*, 1636–1652. <http://dx.doi.org/10.1287/mnsc.49.12.1636.25109>
- Bloom, P. (2010). How do morals change? *Nature*, *464*, 490–490. <http://dx.doi.org/10.1038/464490a>
- Broome, J. (2013). *Rationality through reasoning*. Hoboken, NJ: Wiley. <http://dx.doi.org/10.1002/9781118609088>
- Bucciarelli, M., & Daniele, M. (2015). Reasoning in moral conflicts. *Thinking & Reasoning*, *21*, 265–294. <http://dx.doi.org/10.1080/13546783.2014.970230>
- Bucciarelli, M., Khemlani, S., & Johnson-Laird, P. N. (2008). They psychology of moral reasoning. *Judgment and Decision Making*, *3*, 121–139.
- Coles, R. (1986). *The moral life of children*. Boston, MA: Atlantic Monthly Press.
- Crisp, R. (2005). Value, reasons and the structure of justification: How to avoid passing the buck. *Analysis*, *65*, 80–85. <http://dx.doi.org/10.1093/analysis/65.1.80>
- Dietrich, F., & List, C. (2016). Reason-based choice and context-dependence: An explanatory framework. *Economics and Philosophy*, *32*, 175–229. <http://dx.doi.org/10.1017/S0266267115000474>
- Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584. <http://dx.doi.org/10.1037/0022-3514.63.4.568>
- Ditto, P. H., Munro, G. D., Apanovitch, A. M., Scepansky, J. A., & Lockhart, L. K. (2003). Spontaneous skepticism: The interplay of motivation and expectation in responses to favorable and unfavorable medical diagnoses. *Personality and Social Psychology Bulletin*, *29*, 1120–1132. <http://dx.doi.org/10.1177/0146167203254536>
- Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral

- reasoning. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 50, pp. 307–338). Amsterdam, the Netherlands: Elsevier. [http://dx.doi.org/10.1016/S0079-7421\(08\)00410-6](http://dx.doi.org/10.1016/S0079-7421(08)00410-6)
- Ditto, P. H., Scepanky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, *75*, 53–69. <http://dx.doi.org/10.1037/0022-3514.75.1.53>
- Evers, D. (2010). The end-relational theory of 'ought' and the weight of reasons. *Dialectica*, *64*, 405–417. <http://dx.doi.org/10.1111/j.1746-8361.2010.01235.x>
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Goodman, L. A. (1969). How to ransack social mobility tables and other kinds of cross-classification tables. *American Journal of Sociology*, *75*, 1–40. <http://dx.doi.org/10.1086/224743>
- Greene, J. D. (2007). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 3: The neuroscience of morality: Emotion, disease, and development*. Cambridge, MA: MIT Press.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*, 517–523. [http://dx.doi.org/10.1016/S1364-6613\(02\)02111-9](http://dx.doi.org/10.1016/S1364-6613(02)02111-9)
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144–1154. <http://dx.doi.org/10.1016/j.cognition.2007.11.004>
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834. <http://dx.doi.org/10.1037/0033-295X.108.4.814>
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002. <http://dx.doi.org/10.1126/science.1137651>
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Vintage.
- Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Lund Psychological Reports*, *1*.
- Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*, 191–221. <http://dx.doi.org/10.1111/j.1559-1816.2001.tb02489.x>
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*, 613–628. <http://dx.doi.org/10.1037/0022-3514.65.4.613>
- Horne, Z., Powell, D., & Hummel, J. (2015). A single counterexample leads to moral belief revision. *Cognitive Science*, *39*, 1950–1964. <http://dx.doi.org/10.1111/cogs.12223>
- Hume, D. (1965). *An enquiry concerning the principles of morals*. Notre Dame, IN: University of Notre Dame Press. (Original work published 1777)
- Kearns, S., & Star, D. (2009). Reasons as evidence. *Oxford Studies in Metaethics*, *4*, 215–242.
- Kearns, S., & Star, D. (2013). Weighing reasons. *Journal of Moral Philosophy*, *10*, 70–86. <http://dx.doi.org/10.1163/174552412X628878>
- Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences of the USA*, *113*, 6166–6171.
- Kuhn, D. (1991). *The skills of argument*. Cambridge, MA: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571350>
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498. <http://dx.doi.org/10.1037/0033-2909.108.3.480>
- Lieberman, A., & Schroeder, M. (2016). Commitment: Worth the weight. In E. Lord & B. Maguire (Eds.), *Weighing Reasons* (pp. 104–120). Oxford, UK: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199315192.003.0005>
- Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., & Blair, R. J. R. (2006). The neural basis of implicit moral attitude—An IAT study using event-related fMRI. *NeuroImage*, *30*, 1449–1457. <http://dx.doi.org/10.1016/j.neuroimage.2005.11.005>
- Maki, A., & Raimi, K. T. (2017). Environmental peer persuasion: How moral exporting and belief superiority relate to efforts to influence others. *Journal of Environmental Psychology*, *49*, 18–29. <http://dx.doi.org/10.1016/j.jenvp.2016.11.005>
- Morgan, G. S., Skitka, L. J., & Wisneski, D. C. (2010). Moral and religious convictions and intentions to vote in the 2008 presidential election. *Analyses of Social Issues and Public Policy*, *10*, 307–320. <http://dx.doi.org/10.1111/j.1530-2415.2010.01204.x>
- Nair, G. S. (2016). How do reasons accrue? In E. Lord & B. Maguire (Eds.), *Weighing reasons* (pp. 56–73). Oxford, UK: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199315192.003.0003>
- Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, *87*, 817–831. <http://dx.doi.org/10.1037/0022-3514.87.6.817>
- Parfit, D. (2011). *On what matters*. Oxford, UK: Oxford University Press.
- Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in Cognitive Science*, *2*, 511–527. <http://dx.doi.org/10.1111/j.1756-8765.2010.01096.x>
- Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, *36*, 163–177. <http://dx.doi.org/10.1111/j.1551-6709.2011.01210.x>
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 123–205). New York, NY: Academic Press.
- Philips, M. (1987). Weighing moral reasons. *Mind*, *96*, 367–375. <http://dx.doi.org/10.1093/mind/XCVI.383.367>
- Pinillos, N. Á., Smith, N., Nair, G. S., Marchetto, P., & Mun, C. (2011). Philosophy's new challenge: Experiments and intentional action. *Mind & Language*, *26*, 115–139. <http://dx.doi.org/10.1111/j.1468-0017.2010.01412.x>
- Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: Comment on Haidt (2001). *Psychological Review*, *110*, 193–196. <http://dx.doi.org/10.1037/0033-295X.110.1.193>
- Pyszczynski, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. *Advances in Experimental Social Psychology*, *20*, 297–340. [http://dx.doi.org/10.1016/S0065-2601\(08\)60417-7](http://dx.doi.org/10.1016/S0065-2601(08)60417-7)
- Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179. <http://dx.doi.org/10.1016/j.jtbi.2011.03.004>
- Ross, W. D. (1930). *The right and the good*. Oxford, UK: Clarendon Press.
- Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, *10*, 296.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*, 1755–1758. <http://dx.doi.org/10.1126/science.1082976>
- Scanlon, T. M. (1998). *On what we owe to each other*. Boston, MA: Belknap Press.
- Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11–36. [http://dx.doi.org/10.1016/0010-0277\(93\)90034-S](http://dx.doi.org/10.1016/0010-0277(93)90034-S)
- Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*, 158–174. <http://dx.doi.org/10.1086/209205>
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, *4*, 267–281. <http://dx.doi.org/10.1111/j.1751-9004.2010.00254.x>

- Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology, 88*, 895–917. <http://dx.doi.org/10.1037/0022-3514.88.6.895>
- Snedegar, J. (2017). Reasons for and reasons against. *Philosophical Studies*. Advance online publication. <http://dx.doi.org/10.1007/s11098-017-0889-2>
- Stanley, M. L., Henne, P., Iyengar, V., Sinnott-Armstrong, W., & De Brigard, F. (2017). I'm not the person I used to be: The self and autobiographical memories of immoral actions. *Journal of Experimental Psychology: General, 146*, 884–895. <http://dx.doi.org/10.1037/xge0000317>
- Strohinger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131*, 159–171. <http://dx.doi.org/10.1016/j.cognition.2013.12.005>
- Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511613500>
- Ugazio, G., Lamm, C., & Singer, T. (2012). The role of emotions for moral judgments depends on the type of emotion and moral scenario. *Emotion, 12*, 579–590. <http://dx.doi.org/10.1037/a0024611>
- Uhlmann, E., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science, 16*, 474–480.
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science, 17*, 476–477. <http://dx.doi.org/10.1111/j.1467-9280.2006.01731.x>
- Vlaev, I., Chater, N., Lewis, R., & Davies, G. (2009). Reason-based judgments: Using reasons to decouple perceived price–quality correlation. *Journal of Economic Psychology, 30*, 721–731. <http://dx.doi.org/10.1016/j.joep.2009.06.003>
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*, 780–784. <http://dx.doi.org/10.1111/j.1467-9280.2005.01614.x>
- Wiech, K., Kahane, G., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2013). Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. *Cognition, 126*, 364–372. <http://dx.doi.org/10.1016/j.cognition.2012.11.002>
- Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. *Advances in Experimental Social Psychology, 22*, 287–343. [http://dx.doi.org/10.1016/S0065-2601\(08\)60311-1](http://dx.doi.org/10.1016/S0065-2601(08)60311-1)
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175. <http://dx.doi.org/10.1037/0003-066X.35.2.151>

## Appendix A

### Philosophical Moral Dilemmas

The following are the dilemmas used in both Study 1 and Study 1b with the respective reasons for each available choice.

#### Crying Baby Dilemma

Enemy soldiers have taken over your village. The soldiers have orders to kill all remaining adult civilians, but no babies or young children. You and some of your townspeople have sought refuge in the cellar of a large house. Outside you hear the voices of soldiers who have come to search the house for valuables. Your baby begins to cry loudly. You cover the baby's mouth to block the sound, but the baby's crying cannot be stopped. If the baby continues to cry, he will summon the attention of the soldiers, who will undoubtedly kill you and the other adults hiding out in the cellar. Smothering the baby to death is the only way to silence him. To save yourself and the others, you would need to smother the baby to death. You have two possible courses of action: You can

choose to smother your baby, or you can choose not to smother your baby.

a. The six reasons provided for smothering the baby:

1. I would save a greater number of lives by smothering the baby than by letting the baby cry.
2. I have an obligation to my fellow refugees to do everything in my power to keep them safe.
3. Smothering the baby is the only way I will be able to survive.
4. Doing so would allow me to avoid feeling the pain and guilt of allowing my fellow refugees to die.
5. Doing so would make it possible for me and my fellow refugees to later relocate to a safer, new home.

(Appendices continue)

6. My friends and family might be angry and disappointed with me for letting the other refugees die when I could have saved them.
- b. The six reasons provided for not smothering the baby:
1. One should not ever harm another human being.
  2. Otherwise I would have to live with the guilt of killing my own child for the rest of my life.
  3. As a parent it is my duty to do everything in my power to protect my own child.
  4. The baby could grow up to do incredible good in the world.
  5. There is a chance that the soldiers might not follow orders after feeling sympathy toward a helpless infant and parent.
  6. My own friends and family might be angry and disappointed with me for killing my child.
4. My son was aware of the consequences of getting caught trying to escape, yet he was willing to take the risk.
  5. I would be with my son, serving as a source of comfort, in the last moments of his life.
  6. My friends and family might be angry and disappointed with me for letting another innocent inmate die.
- b. The six reasons provided for not pulling the chair out from underneath the son:
1. The act of rebelling might inspire other inmates to revolt and gain freedom.
  2. If both my son and some other inmate die, there might be more food for the remaining inmates.
  3. It is always morally wrong to harm another human being.
  4. As a parent, I have an obligation to not cause any harm to my own child.
  5. If I were to pull the chair, then I would have to live with the guilt of killing my son for the rest of my life.
  6. My friends and family might be angry and disappointed with me for killing my son.

### Concentration Camp Dilemma

You and your son are inmates in a concentration camp. A sadistic guard is about to hang your son who tried unsuccessfully to escape. After putting the noose around your son's neck, the guard wants you to pull the chair from underneath your son. Pulling out the chair would hang your son and surely kill him. If you do not pull out the chair from under your son, the guard said that he would not only kill your son but also some other innocent inmate in the concentration camp as well. You do not have any doubt that the guard means what he says, and you cannot overpower the guard or escape. You have two possible courses of action: You can choose to pull the chair out from underneath your son, or you can choose NOT to pull the chair out from underneath your son.

- a. The six reasons provided for pulling the chair out from underneath the son:
1. Pulling the chair out will prevent the death of another innocent inmate.
  2. The act of obedience may cause the guard to be less harsh in the future.
  3. The concentration camp is a miserable place to be imprisoned, and my son would be better off dead.

### Organ Transplant Dilemma

You are a surgeon working the night shift in an emergency room. A family of five have been in a serious car accident, and each needs an organ transplant immediately. Each member of the family needs a different particular organ transplanted or else each one will surely die. You are currently operating upon an anesthetized patient who suffered a less severe accident. This patient is homeless and has no family. The homeless patient is, however, sufficiently healthy otherwise and would be an ideal organ donor for all members of the family who were in the car accident. If you transplant the homeless patient's organs (against his will) into the bodies of the other patients, all five members of the family will live but the patient will die. But this is the only way to save the family of five. You have two possible courses of action: You can choose to transplant the homeless patient's organs into the bodies of the other five patients, or you can choose NOT to transplant the homeless patient's organs into the bodies of the other five patients.

*(Appendices continue)*

- a. The six reasons provided for transplanting the homeless patient's organs into the bodies of the other five patients:
1. Transplanting the organs would save the greatest number of lives.
  2. Many more friends and family members would be negatively affected by the deaths of the family than the death of the homeless person.
  3. Society would likely benefit from one less homeless person to support.
  4. Otherwise I might feel very guilty for allowing five people to die when I could have done something about it.
  5. The children in the family could grow up to do much more good in the world than the homeless man.
  6. My friends and family might be angry and disappointed with me for letting 5 people die when I could have saved them.
- b. The six reasons provided for not transplanting the homeless patient's organs into the bodies of the other five patients:
1. The homeless man did not provide consent to allow his organs to be harvested.
  2. My medical license would be revoked if I were to get caught.
  3. I would be harming another human being, which is always wrong.
  4. I would be violating the Hippocratic oath if I transplant the organs.
  5. The family members might feel guilty for the rest of their lives for having someone else's life taken to save their own.
  6. My friends and family might be angry and disappointed with me for killing the homeless person to transplant the organs.

## Appendix B

### Everyday Moral Dilemmas

The following are the dilemmas used in study 2a and 2b with the respective reasons for each available choice.

#### **Car Accident Dilemma**

It is late at night, and as you are parking your car, you hear a scraping sound. You are not certain, but you think that you accidentally scratched the car parked next to you as you were pulling into the parking spot. When you get out of your car to inspect the possible damage, you find a fairly minor scratch on the side of the other car. No airbags deployed in either car, and no car alarms went off. Looking around, you notice that nobody else is in the parking lot. You consider leaving a message with your contact information offering to pay for the damage on the windshield of the other damaged car. You have two possible courses of action: you can leave your contact information on a note placed on the windshield of the other damaged car, or you can just drive off without leaving the note.

- a. The seven reasons provided for not leaving contact information on a note offering to pay for the damages:
1. I am pretty confident that nobody saw me, so it is unlikely that there would be negative consequences.
  2. I would be able to use the money that would have gone toward repairing the damages in a way I think is more useful.
  3. The scratch is rather minor, so there is a chance that the owner of the other car will not even notice.
  4. There is a chance that the scratch was already on the car before I hit it, meaning that I could be paying for damages I did not cause.
  5. The scratch is so minor that it will not adversely affect the other driver's ability to drive his or her car.

*(Appendices continue)*

6. The owner of the other car might try to take advantage of the situation and extort me for an unfair amount of money.
  7. Reporting the accident might cause my insurance premium to increase.
- b. The seven reasons provided for leaving contact information on a note offering to pay for damages:
1. I would be responsible for damaging another person's property and should accept responsibility for causing the damage.
  2. The guilt I might feel over not leaving a note might outweigh the financial benefit of not having to pay to fix the scratch.
  3. By having to pay for the damages, I will be discouraged from making a similar mistake in the future and become a more careful driver.
  4. The owner of the other car might be more seriously affected by the financial loss that comes with the car scratch than I am by paying for the scratch.
  5. I will feel good about myself for offering to pay for the damages to the other car.
  6. Although no one else was seen in the parking lot, there is always a chance that someone saw and reported me.
  7. My friends and family might be disappointed with me for not taking responsibility for my actions.

### **Cheating Dilemma (adapted from Kouchaki & Gino, 2016)**

You have to take one general requirement chemistry class in your final semester of college to get your bachelor's degree in anthropology. The professor teaching the chemistry class is a really poor teacher who has given unnecessarily challenging exams and has graded harshly. When finals come around, you are close to failing the class. You study very hard and even pulled an "all nighter" the night before the final exam. However, you do not feel as though you are retaining any of the information. As a result, you made a cheat sheet as a backup even though it was not allowed. You could not remember the answers to several questions on the final exam, and you consider using the cheat sheet. You even see three other students cheating on the exam. You have two possible courses of action: you can use the cheat sheet on the exam, or you cannot use the cheat sheet on the exam.

- a. The seven reasons provided for using the cheat sheet on the exam:

1. The professor was a really poor teacher.
  2. If I fail the final, I will have to retake the class and not graduate on time.
  3. I will not be using anything I learned in this chemistry class later in life anyway.
  4. Other students in the class were cheating, so not cheating would put me at a disadvantage.
  5. The professor was an unnecessarily harsh grader.
  6. Doing poorly in the class might make it more difficult for me to get a job after graduation.
  7. My parents would likely be disappointed in me if I fail the class.
- b. The seven reasons provided for not using the cheat sheet on the exam:
1. It is wrong to cheat on anything whatsoever.
  2. If I am willing to cheat once, I might be more likely to cheat again in the future.
  3. If anyone notices me cheating, it might harm my reputation.
  4. If I am caught cheating, my parents would likely be very disappointed.
  5. Cheating on the exam might reduce the enjoyment of my accomplishments earned through genuine effort.
  6. Cheating on the exam would create an environment of broken trust limiting the ability for students and faculty to collaborate meaningfully.
  7. If I am caught cheating by the professor, there will be serious consequences.

### **Extra Change Dilemma**

You were almost finished with your weekly trip to the grocery store. The grocery store you are shopping at is part of a large chain of grocery stores with locations throughout the United States. You had been waiting in line at the grocery store to pay for your groceries for quite some time. Once you finally reach the front of the line to check out, the clerk ignores you for a moment while she sends a text on her phone. Once she finishes sending her text, she checks you out. You pay with cash. The clerk accidentally gives you an extra \$40 in change. You just lost your job, so the extra \$40

*(Appendices continue)*

would be helpful while you search for a new job. You have two possible courses of action: You can return the money to the clerk, or you can just walk out of the grocery store with the extra money.

- a. The seven reasons provided for returning the extra money to the clerk:
  1. If the manager learns about the missing money, the clerk might lose her job.
  2. The missing \$40 may be taken out of the clerk's pay-check.
  3. If I were in the position of the clerk, I would want customers to return the money.
  4. Being honest is always the best course of action.
  5. You might feel particularly guilty for keeping the extra money.
  6. The clerk could have given me the extra money intentionally to test my character.
  7. My friends and family might be disappointed in me for not returning the money.
- b. The seven reasons provided for not returning the extra money to the clerk:
  1. This will teach the grocery store clerk to be more responsible.
  2. It is not my fault that the clerk gave you extra change.
  3. It may help reveal to the store administrators that the clerk is an unprofessional employee.
  4. The clerk was being rude by texting on the job anyways.
  5. The extra \$40 will help me make ends meet while I search for a new job.
  6. The grocery store is part of a large corporation that would not suffer from losing \$40.
  7. The extra money will make up for the inconvenience of waiting in line for so long.

Received May 9, 2017

Revision received August 8, 2017

Accepted August 10, 2017 ■