Judgment and Decision Making, Vol. 12, No. 5, September 2017, pp. 481–490

Decisions in moral dilemmas: The influence of subjective beliefs in outcome probabilities

Yiyun Shou* Fei Song†

Abstract

Previous studies have found that the proportions of people who endorsed utilitarian decisions varied across different variants of the trolley dilemma. In this paper, we explored whether moral choices were associated with beliefs about outcome probabilities in different moral dilemmas. Results of two experiments showed that participants' perceptions of outcome probabilities were different between two dilemmas that were similar to the classical *switch* case and *footbridge* case. Participants' judgments of the outcome probabilities were associated with their moral choices. The results suggested that participants might not accept task instructions and thus did not perceive the outcomes in the dilemmas as certain. We argued that researchers who endorse descriptive tasks in moral reasoning research should be cautious about the findings and should take participants' beliefs in the outcomes into account.

Keywords: moral decision, probability judgment, moral reasoning, moral dilemma

1 Introduction

Moral reasoning has been under long-term intellectual scrutiny. Recent psychological investigations of moral reasoning frequently employ moral dilemmas (Crockett, 2013; Cushman, Young & Hauser, 2006; Greene et al., 2009; Greene, Morelli, Lowenberg, Nystrom & Cohen, 2008; Haidt, 2007; Moll & de Oliveira-Souza, 2007). The well-known trolley dilemma requires people to choose between two options: killing one person to save five or letting the five people die. Moral dilemmas such as this commonly engender conflict between two major approaches to moral reasoning: consequentialist and deontological approaches. The consequentialist approach primarily concerns the outcome of each option and aims at choosing the one with the best outcome. By contrast, the deontological approach is concerned with whether an act is consistent with a moral principle or duty. In most studies, the choice of killing (directly or indirectly) one person in the trolley dilemma is taken as a result of maximizing the outcome utility (i.e., the number of people who will not die) and is often associated with the consequentialist approach. Not killing, on the other hand, is taken as a product of the deontological approach under which the action of killing is regarded as a deontological violation (Cummins & Cummins, 2012).

Previous studies revealed that people's moral judgments (i.e., the choice of which option is more morally acceptable) varied across different dilemmas (Cummins & Cummins, 2012; Greene, Sommerville, Nystrom, Darley & Cohen, 2001). In the *switch* case, participants have to decide between flipping a switch to shift a trolley to a side-track on which one person is trapped, and doing nothing. The majority of participants perceived the choice of switching the trolley (consequently, killing the minority) as morally preferable. However, most participants preferred doing nothing (or not killing the minority) in other variants of the trolley dilemma, such as the well-known footbridge dilemma (e.g., Crockett, 2013; Greene et al., 2001, 2009; Lerner, Li, Valdesolo & Kassam, 2014; Shenhav & Greene, 2014; Valdesolo & DeSteno, 2006; Youssef et al., 2012). In the footbridge dilemma, one is required to decide whether or not to push a fat man over a footbridge to stop a runaway trolley that would otherwise kill five people. Researchers have asked why there is a discrepancy in moral judgments between these two types of dilemmas.

Many studies have attributed the differences to certain features of moral dilemmas. For example, personal dilemmas (e.g., the footbridge case) involve harm with personal forces and thereby a stronger negative affect would be elicited, which has often been characterized as the main reason underlying the commonality of the preference for not killing the individual (Greene, Nystrom, Engell, Darley & Cohen, 2004). A methodological problem with most studies is that they equate the choice of killing vs. not killing with the endorsement of the consequentialist/deontological reasoning approach, upon the presumption that participants perceived the outcomes in moral dilemmas as certain (Kortenkamp &

*Research School of Psychology, The Australian National University, Canberra, ACT 2601, Australia. Email: yiyun.shou@anu.edu.au..

†Department of Philosophy, The University of Hong Kong, Hong Kong.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

Moore, 2014). However, participants may hold or be influenced by subjective beliefs regarding how likely an outcome is to occur when a particular choice is taken. These beliefs consequently influence their expected utilities of the given choices in different moral dilemmas. The perceived uncertainty in the outcomes could reduce the conflict between deontological versus consequentialist approaches if the action of killing is no longer producing the best outcomes.

In this paper, we explored whether participants' choice preferences are associated with subjective beliefs about the probabilities of the outcomes in the dilemmas. Choosing killing (hereafter K) or not killing (hereafter $\tilde{K}$) as the morally preferred choice does not necessarily reflect deontological or consequentialist approaches. Though $\tilde{K}$ is usually taken as a result of deontological reasoning, a consequentialist may also choose $\tilde{K}$ when the expected value of K is lower than that of $\tilde{K}$. For instance, in the footbridge dilemma, one may assign a probability lower than 100% to the outcome that sees the trolley being stopped by the fat man. In this case, the aversion to "doing harm" in the footbridge dilemma could be taken either as a result of deontological reasoning or as a consequence of the aversion to uncertainty (Rogers, Viding & Chamorro-Premuzic, 2013).

2 Study 1

Study 1 examined how participants perceived the outcome probabilities in different types of dilemmas and how their judgments of outcome probabilities were associated with their moral choices. We utilized two dilemmas that were generated by Shou and Song (2014). One involved personal force while the other did not (see materials for details). In accordance with dilemmas used in previous studies, the dilemmas in the current study specified the positive outcome (or benefit) and the negative outcome for each of the two choices (K or $\tilde{K}$). The positive outcome of K was the survival of five people, while its negative outcome was the death of one person. On the other hand, the positive outcome of $\tilde{K}$ was the survival of one person, while its negative outcome was the death of five people. In each dilemma, participants provided probability judgments for the positive and negative outcomes from a given choice (one for each of the four parts of a dilemma, see below for more details).

We hypothesized that:

H1: Participants' judgments of the outcome probabilities would be differ between the two dilemmas.

H2: Participants' moral choices would be associated with their perceived outcome probabilities. Higher probabilities of positive outcomes and lower probabilities of negative outcomes given a particular choice would be associated with a higher likelihood of endorsing that choice.

2.1 Method

2.1.1 Participants and procedure

A total of 112 participants (85 females) were recruited via online crowd-sourcing service, CrowdFlower. Their ages were between 19 and 71, with a mean of 40.9 years (SD = 11.67). Participants were randomly assigned to one of two moral dilemma scenarios (described below). They read the consent information and completed the demographical questionnaire and the moral judgment task in order.

2.1.2 Materials

One of the two vignettes described a *car* dilemma, where the participant was asked to imagine that s/he was driving a truck approaching a sharp turn near a cliff. A car of five passengers suddenly stopped in front of the truck. The participant had two options, of which the K choice was to turn the truck into one bystander and the other option was to let the truck hit the car with the five people inside. The other vignette described a *hostage* dilemma, in which the participant was passing by a cliff with another innocent person. The participant was threatened by a gangster who had captured five hostages. The participant had two options: to push the other person over the cliff or to let the gangster shoot the five hostages.[1] The car and hostage dilemmas differed in terms of whether the agent (the participant) had physical contact with the victim. This personal/impersonal distinction was similar to one of the differences between the trolley and the footbridge dilemmas.

The description of the two dilemmas did not contain any probabilistic information, but indicated that the individual victim (passerby or the innocent person) would die if K was chosen, while the five victims (five passengers or hostages) would die if $\tilde{K}$ was chosen. After reading the assigned dilemma, participants were asked, "Which action do you think is morally better?" We used this question to focus participants on the issue of morality. Baron (2013) argued that conventional moral judgment questions that use terms such as "permissible" may draw participants' attentions to law or convention and away from morality. In addition, we avoided asking questions regarding which action the participant would choose to distinguish moral judgment from preference.

After making the moral choice, participants provided probability judgments for four possible outcomes. As mentioned earlier, the positive outcome (PO) given K is the survival of five people, while its negative outcome (NO) is the death of the individual. The positive outcome given $\tilde{K}$ is the survival of the individual, while its negative outcome is the death of the five people. Participants provided judgments for the following probabilities:

[1]The details of materials as well as example illustrations are available in the supplement.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

1. $P(\text{PO}|K)$: the probability that five people would survive given K is chosen;
2. $P(\text{NO}|K)$: the probability that one person would die given K is chosen;
3. $P(\text{NO}|\tilde{K})$: the probability that five people would die given $\tilde{K}$ is chosen;
4. $P(\text{PO}|\tilde{K})$: the probability that one person would survive given $\tilde{K}$ is chosen.

2.2 Results

For the moral choice, participants in the car dilemma were significantly more likely to choose K than those in the hostage dilemma, $\chi^2(1) = 13.76$, $p < .001$. About 82% of the participants (47 out of 57) in the car dilemma indicated that killing the individual was morally better than letting five people die, compared to 47% of the participants (26 out of 55) in the hostage dilemma. Figure 1 shows the means of participants' probability judgments in two dilemmas.

We first tested H1, that participants' probability judgments would be different between different dilemmas, using independent-sample t-tests. Participants in the car dilemma provided significantly higher P(PO|K) and P(PO|$\tilde{K}$) than the participants in the hostage dilemma did, $t(110) = 5.22$ and 4.00, respectively, ps <. 001. Participants' estimates for P(NO|K) and P(NO|$\tilde{K}$) did not significantly differ between the two dilemmas, $t(110) = -1.49$ and 1.07, $p = .139$ and .287, respectively. We then compared the probability judgments between the two given choices using paired t-tests. The mean perceived positive outcome probabilities were not significantly different between K and $\tilde{K}$ in either the car or the hostage dilemmas. $|t| < 1$, ps> .370. The mean perceived negative outcome probabilities were also not significantly different between K and $\tilde{K}$ in the car dilemma, $t(56) = -0.27$, $p = .786$. However, the perceived negative outcome probabilities given K were significantly higher than those given $\tilde{K}$ in the hostage dilemma, $t(54) = 2.70$, $p = .009$.

2.2.1 Association between moral choices and probability judgments

Logistic regression was used to examine the associations between participants' moral choices and their probability judgments. The four probability judgments and dilemma types were used to predict participants' moral choices. Likelihood ratio tests suggest that there were no significant interaction effects between dilemma type and any of the four probability judgments, ps > .059.

Table 1 shows the estimation results of the final model. Participants were more likely to choose K when their perceived probability of the positive outcome given K ($P(\text{PO}|K)$) was higher, $b = 1.28$, $p = .002$, or when their perceived probability of the negative outcome given $\tilde{K}$ ($P(\text{NO}|\tilde{K})$) was higher, $b = 1.56$, $p < .001$. The perceived positive outcome given $\tilde{K}$ ($P(\text{PO}|\tilde{K})$) reduced the likelihood of choosing K, while the perceived negative outcome given K ($P(\text{NO}|K)$) did not have a significant influence on participants' moral decisions.

Table 1: Logistic regression model for moral choices by the types of dilemma, the four probability judgments and order of task.

	r	*b*	*SE*	*t*	*p*
Intercept		1.31	0.37	3.54	<.001
$P(\text{NO}\|K)$	0.00	–0.17	0.28	–0.61	.541
$P(\text{PO}\|K)$	0.36	1.28	0.42	3.04	.002
$P(\text{NO}\|\tilde{K})$	0.48	1.56	0.37	4.23	<. 001
$P(\text{PO}\|\tilde{K})$	0.00	–0.71	0.32	-2.23	.026
Dilemma (car)		1.01	0.33	3.07	.002

Note. $P(\text{PO}|K)$: the probability that five people would survive given that K is chosen; $P(\text{NO}|K)$: the probability that one person would die given that K is chosen; $P(\text{NO}|\tilde{K})$: the probability that five people would die given that $\tilde{K}$ is chosen; $P(\text{PO}|\tilde{K})$: the probability that one person would survive given that $\tilde{K}$ is chosen. The four probabilities were standardized. r is the Pearson's correlations between the probability estimates and the dummy coded choices.

2.3 Discussion

Study 1 examined how participants perceived the outcome probabilities in different dilemmas. Overall, participants in the car dilemma provided probability judgments for the positive outcomes that were significantly higher than those in the hostage dilemma. With regards to the negative outcomes, participants in the hostage dilemma perceived a higher likelihood that the individual would die if K (killing one person) had been chosen, than that the five people would die if $\tilde{K}$ (not killing) had been chosen. It is clear that participants perceived the expected utility given K in the car dilemma to be higher than for the hostage dilemma. This may explain why there was a higher proportion of participants in the car dilemma that chose K than in the hostage dilemma.

We also examined the association between participants' probability judgments and their moral decisions. The probabilities of outcomes that involved the five people had a greater association with participants' moral choices for both the car and hostage dilemmas than those that involved the individual. Participants were more likely to choose K if they perceived the five people had a higher chance of surviving ($P(\text{PO}|K)$) with the sacrifice of the individual, or that the five people had a higher chance of dying if they did nothing. These results suggested that participants preferred $\tilde{K}$ more

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

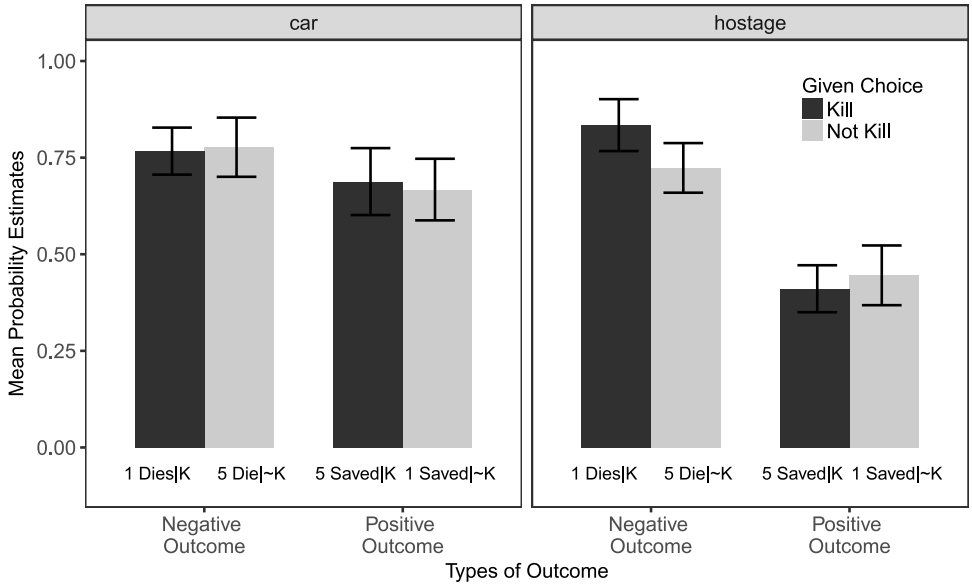

FIGURE 1: Mean probability judgments for the outcomes given the two alternatives between participants in the car dilemma and those in the hostage dilemma. (Error bars are standard errors.)

in the hostage dilemma than in the car dilemma, which might be due to a perception of higher loss and lower gain of K in the car dilemma. This also implied that participants were likely applying consequentialist reasoning even though they chose $\tilde{K}$.

Two confounding factors in Study 1, however, raise questions. First, the probability judgments were assessed after participants had made the moral choices. Participants may have provided the judgments to better justify their choices. Second, descriptions of the two dilemmas explicitly stated only that the negative outcomes would happen for each choice. The low probability judgments of the positive outcomes might be due to the lack of explicit instruction. These two issues are addressed in Study 2.

3 Study 2

To examine whether participants in Study 1 provided probability judgments to justify choices they have made, Study 2 included three conditions. The first condition was the same as Study 1, where participants made moral choices prior to their probability judgments. In the second condition, participants provided probability judgments first and then made moral choices. In a third condition, instead of being forced to make a choice from the two alternatives, participants were provided with a third option that would result in no death in the scenario.[2] The third option would be obviously the best option and we expected most participants would take that option. Thus, participants would not need to justify their choices when providing the outcome probability judgments for the two original options. We hypothesized that participants' probability estimates would not significantly differ across the three conditions.

In addition, to address the second issue of Study 1, both the negative (one or five people will die) and positive (one or five "will not die") outcomes were explicitly described in the task instruction. If participants ignore or refuse to accept information (i.e., "will die" statements) in the task instruction, the pattern of probability judgments in Study 2 would be similar to that in Study 1 (i.e., higher probability judgments for the negative outcomes than for the positive outcomes).

3.1 Method

3.1.1 Participants and procedure

A total of 401 participants (282 females) were recruited via the online crowd-sourcing service, CrowdFlower. Their ages were between 18 and 67, with a mean of 38.69 years (SD = 12.61). The study had a 3 (conditions) by 2 (moral dilemmas) between-subject design. Participants were randomly assigned to one of the six conditions. The three experimental conditions were: (1) 2-alternative with moral choices taken before probability judgments (2Alt-Dfirst); (2) 2-alternative with moral choices taken after probability judgments (2Alt-Dlater); (3) 3-alternative with moral choices taken after probability judgments (3Alt).

3.1.2 Materials

The two vignettes were the same as Study 1, with descriptions for both positive and negative outcomes given a choice.

[2]We are grateful to Jonathan Baron for suggesting this design.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

The description of the two choices for participants in the car dilemma was: *"Turn the truck and hit the passerby, the person will die and the five people will not die. Do nothing, and hit the car in front of you, the car will fall off the cliff and the five people inside will die and the person will not die."* The description in the hostage dilemma was: *"Push James off the cliff, James will die and the five people will not die. Do nothing, the gangster will shoot all of the five people and the five people will die and James will not die."*

In the 2Alt-Dfirst condition, participants provided their moral choices first, and then provided probability judgments for the four types of outcomes outlined in Study 1 in a subsequent survey page. In the 2Alt-Dlater condition, participants made probability judgments first and provided moral choices on the subsequent page.

In the 3Alt condition, participants were given three options, the third of which was an action that would result in no one dying. The third option in the car dilemma was "*Turn the truck right and keep braking, neither the person nor the five people inside the car will die*". The third option in the hostage dilemma was "*Try to negotiate with the gangster to buy some time for police to come, and neither the five hostages nor James will die*." Participants first made the moral decisions, and then provided the four probability judgments on a following page.

3.2 Results

Table 2 summarizes the frequencies and proportions of choices across different conditions. Among the participants in the 3Alt condition, 87% of the participants chose the third option in the car dilemma, while 91% chose the third option in the hostage dilemma. We included participants who chose the third option for the following analyses.[3].

Figure 2 shows the mean probability judgments in the two dilemmas, grouped by the different experimental conditions. The mean probability judgments were similar across different conditions. The pattern of the results was also similar to the one in Study 1.

3.2.1 Effects of case, given choice and condition on probability judgments

We first examined whether participants' probability judgments would be significantly different between the two dilemmas, and between different conditions. In the following analyses, linear regression models were used to analyze the probability judgments. The models included Given Choice, Dilemma Type and the interaction between the two, as the predictor terms. Analyses were carried out using R version 3.1.0.

[3]We also analysed the results with including all participants, the results were similar to the ones that are presented here.

The probability judgments for the negative outcomes and the probability judgments for the positive outcomes were modeled separately. Results of the scope of model comparison are displayed in the Appendix. The probability judgments for the negative outcomes were significantly different between the two choices, $F(1, 774) = 47.72$, $p < .001$. Neither the dilemma type nor condition had a significant main effect on the probability judgments. The effect of the Given Choice was moderated by dilemma type, $F(1, 770) = 8.75$, $p = .003$. No other interactions were found. Results for the positive outcome probabilities are displayed in the second part of the table. The probability judgments of the positive outcomes were significantly different between the two dilemmas, $F(1, 774) = 184.06$, $p < .001$. This effect was significantly moderated by Given Choice, $F(1, 770) = 9.97$, $p = .002$. Condition did not significantly moderate the effects of the other predictors.

Table 3 shows the results of the final models for the positive and negative outcome probability judgments. Participants in the hostage dilemma had significantly higher probability judgments for the negative outcome when K was chosen than when $\tilde{K}$ was chosen, while participants in the car dilemma had similar probability judgments for the negative outcomes given the two choices. Participants in the car dilemma had significantly higher probability judgments for the positive outcomes than those in the hostage dilemma regardless of the given choices.

3.2.2 Association between choices and probability judgments

Logistic regression was carried out to test the association between the four types of probability judgments and participants' moral choices. We used a likelihood ratio test to test the contribution of the interaction effects between different predictors in a similar approach to that utilized in the previous section. There were no significant interaction effects between condition and the four probability judgments or dilemma type, ps > .082.

Table 4 shows the results of the final model. Participants were more likely to choose K when their perceived negative outcome given K ($P(\text{NO}|K)$) was lower, $b = -0.67$, $p < .001$, when their perceived positive outcome given K ($P(\text{PO}|K)$) was higher, $b = 0.79$, $p < .001$, or when their perceived negative outcome given $\tilde{K}$ ($P(\text{NO}|\tilde{K})$) was higher, $b = 0.56$, $p < .001$. The perceived positive outcome probability given $\tilde{K}$ did not have a significant influence on participants' moral decisions. It was also noticed that after controlling for the effects of the four probability judgments, the effect of dilemma type was not significant ($p = .099$). This suggests that the perceptions of outcome probabilities may have substantially accounted for the differences in choices between the car and the hostage dilemmas.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

TABLE 2: Frequencies and proportions of choices across different conditions, Study 2.

	Car			Hostage		
	˜*K*	*K*	3rd Option	˜*K*	*K*	3rd Option
2-Option-Dfirst	22 (30%)	49 (70%)	–	46 (61%)	29 (39%)	–
2-Option-Dlater	15 (23%)	51 (77%)	–	40 (56%)	31 (43%)	–
3Alt	5 (8%)	3 (5%)	53 (87%)	3 (5%)	2 (4%)	52 (91%)

TABLE 3: Linear regression model for the probability judgments predicted by the given choice and types of dilemma.

	b	*SE*	*t*	*p*
DV: Negative Outcome Probability				
Intercept	0.77	0.01	90.74	<.001
Given Choice	0.06	0.01	6.88	<.001
Dilemma	0.00	0.01	–0.09	.931
Condition 1	–0.01	0.01	–1.06	.291
Condition 2	–0.02	0.01	–1.40	.161
Dilemma x Given Choice	–0.03	0.01	–2.96	.003
DV: Positive Outcome Probability				
Intercept	0.54	0.01	51.36	<.001
Given Choice	0.03	0.01	2.99	.003
Dilemma	0.14	0.01	13.71	<.001
Condition 1	–0.01	0.02	–0.52	.606
Condition 2	–0.00	0.01	–0.08	.936
Dilemma x Given Choice	0.03	0.01	3.16	.002

Note. Dummy coding for Given Choice: K = 1 and ˜K = –1; Dummy coding for Dilemma: car = 1 and hostage = –1; Condition 1: 2-Option-Dlater = –1, 3Alt = 1, 2-Option-Dfirst = 0; Condition 2: 2-Option-Dlater = –1, 3Alt = 0, 2-Option-Dfirst = 1.

TABLE 4: Logistic regression model for moral choices by the types of dilemma, the four probability judgments and order of task.

	r	*b*	*SE*	*t*	*p*
Intercept		0.36	0.14	2.58	0.010
P(NO|*K*)	–0.22	–0.67	0.16	–4.10	< .001
P(PO|*K*)	0.38	0.79	0.17	4.56	< .001
P(NO|˜*K*)	0.16	0.56	0.15	3.63	< .001
P(PO|˜*K*)	0.11	–0.23	0.16	–1.44	.151
Dilemma (car)		0.27	0.16	1.65	.099
Condition (Dfirst v Dlater)		–0.14	0.14	–1.03	.305

Note. *P*(PO|*K*): the probability that five people would survive given that *K* is chosen; *P*(NO|*K*): the probability that one person would die given that *K* is chosen; *P*(NO|˜*K*): the probability that five people would die given that ˜*K* is chosen; *P*(PO|˜*K*): the probability that one person would survive given that ˜*K* is chosen.

The associations between the four probability estimates and moral choices were also examined for the two dilemmas separately, and the results are shown in Table 5. The directions of the effects for the four probabilities were the same for the two dilemmas. However, the magnitudes of the associations were much stronger in the hostage dilemma than they were in the car dilemma. This result might be due to a ceiling effect in the car dilemma, as the probability judgments for the car dilemma were more negatively skewed (as they were less distorted) and had less variability than they were in the hostage dilemma.

3.2.3 Association between choices and expected values

Finally, we investigated how expected outcome values of the two choices were related to participants' moral judgments. The expected value (based on the number of lived can survive in each outcome) of *K* was calculated as: EV(*K*) = 5**P*(PO|*K*) – 1*(1 – *P*(NO|*K*)); while the expected value of ˜*K* was calculated as: EV(˜*K*) = 1**P*(PO|˜*K*) – 5*(1 – *P*(NO|˜*K*)). We carried out logistic regression using the same procedure in the previous section, by replacing the four probabilities with EV(*K*) and EV(˜*K*). There were no significant interaction effects between dilemma type, the order of the probability judgments and the two expected value variables: a full factorial model did not have significantly better model fit than a main effect only model, χ^2 = 7.59, *p* = .473. For the final model with EV(*K*), EV(˜*K*) and dilemma type as predictors of the moral choices, EV(*K*) had a significant and positive association with the likelihood of choosing *K*, *b* = 0.74, *p* <.001. On the other hand, EV(˜*K*) had a significant negative association with the likelihood of choosing *K*, *b* = –0.42, *p* <.001.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

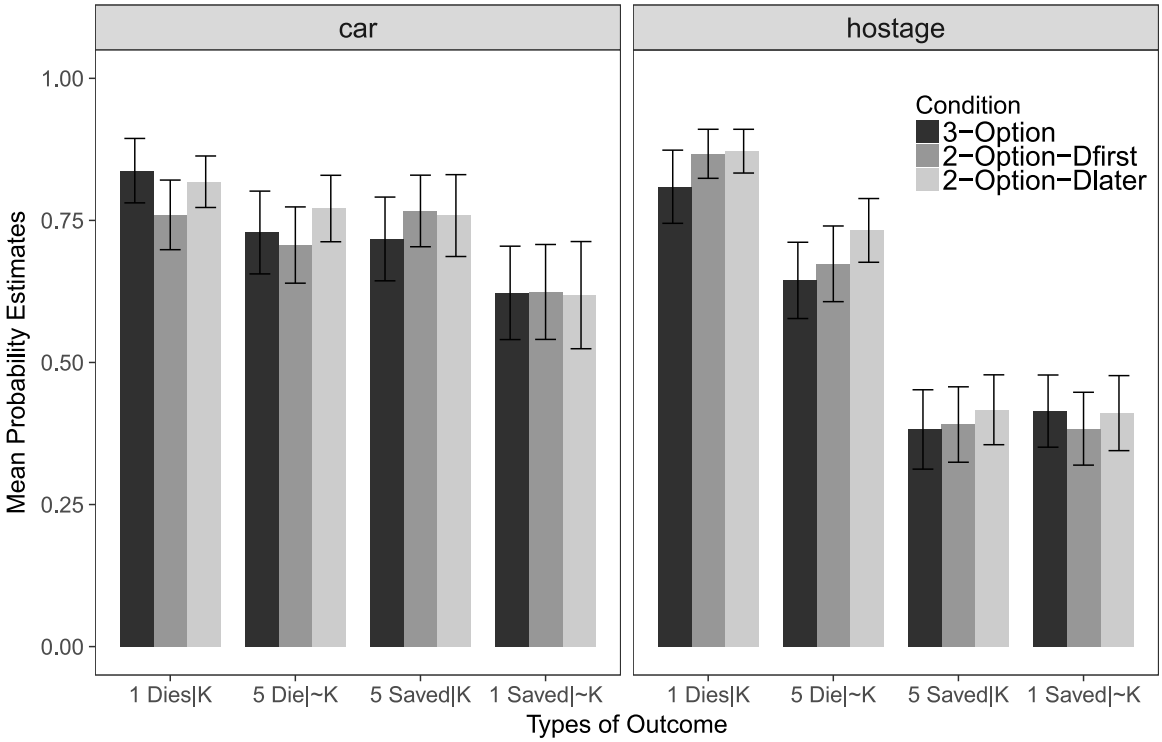

FIGURE 2: Participants' estimated probability in two dilemmas, grouped by the order of probability judgments and making choices. "1 Dies | K" is the probability of the negative outcome given K; "5 Saved | K" is the probability of the positive outcome given K; "5 Die |˜K" is the probability of the negative outcome given $\tilde{}K$; "1 Saved | ˜K is the probability of the positive outcome given $\tilde{}K$.

Figure 3 shows the differences between the EV(K) and EV($\tilde{}K$) for participants choosing K or $\tilde{}K$ across the two dilemmas. Note that almost all participants had the EV(K) higher than the EV($\tilde{}K$) even when their actual choice was $\tilde{}K$. This result implies that perceived outcome probabilities can influence but not fully explain participants' moral judgments.

3.3 Summary and discussion

As expected, the results showed that the patterns of outcome probability judgments were similar across the three conditions. This result provides evidence that the current findings about participants' probability judgments are not due to self-justification in the most obvious way. In addition, similar to Study 1, participants provided lower probability judgments for the positive outcomes than for the negative outcomes, despite the assurance in the task instruction of the positive outcomes. The associations between probability judgments and moral choices were also not significantly different between the condition where choices were taken before probability judgments, and the condition where choices were taken after probability judgments. It was also found that those associations were stronger in the hostage dilemma than in the car dilemma.

4 General discussion

The present paper explored whether the discrepancy in participants' moral choices between two different dilemmas is related to their perceptions of the outcome probabilities. It was found that participants' perceptions of outcome probabilities were significantly different between the two dilemmas utilized in the current study. Participants perceived that the positive outcome was less likely to occur in the hostage dilemma than in the car dilemma. This result also implies that participants perceived that the expected utility of each of the two choices (kill versus not kill) can differ across different dilemmas due to different perceptions of the outcome probabilities.

Participants' perceptions of the outcome probabilities also significantly predicted their moral choices. Participants were less likely to choose a choice if they perceived higher probability of the negative outcome given that choice. The pattern is consistent with a consequentialist approach to moral judgments, where participants prefer a choice that minimizes negative outcomes. Furthermore, this tendency (i.e., avoiding a choice when the negative outcome was more likely) was similar between the two dilemmas in the present study. This suggests that the tendency to endorse a consequentialist approach among participants may not depend on the features

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

TABLE 5: Logistic regression model for moral choices by the four probability judgments in the two dilemmas.

	b	*SE*	*t*	*p*
Car Dilemma				
Intercept	1.10	0.21	5.23	< .001
$P(NO\|K)$	–0.59	0.25	–2.37	.018
$P(PO\|K)$	0.48	0.20	2.39	.017
$P(NO\|\tilde{K})$	0.36	0.22	1.60	.109
$P(PO\|\tilde{K})$	–0.11	0.22	–0.51	.604
Hostage Dilemma				
Intercept	–0.45	0.19	–2.31	0.021
$P(NO\|K)$	–0.64	0.21	–3.12	0.002
$P(PO\|K)$	0.89	0.23	3.91	< .001
$P(NO\|\tilde{K})$	0.78	0.23	3.41	0.001
$P(PO\|\tilde{K})$	–0.27	0.22	–1.26	0.207

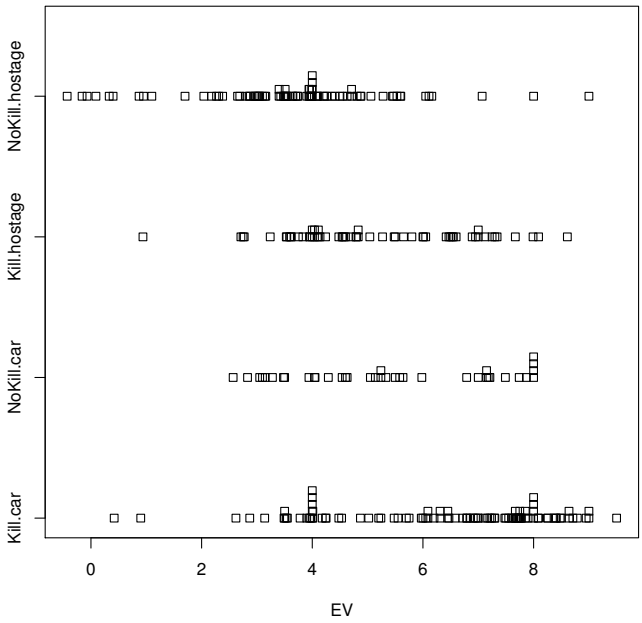

FIGURE 3: The differences between EV(K) and EV($\tilde{K}$) for participants choosing K or ˜K across the two dilemmas.

of the dilemma such as being personal or impersonal, but, rather, on whether the scenario is believable.

Perhaps the most interesting finding was that the mean probabilities of the four outcomes were all well below 100%, even though the instruction stated that these outcomes "*will/will not*" occur given an action. Participants seemed to refuse to believe that one or five individual(s) will or will not die if $\tilde{K}$ or K was taken. This hinges on a phenomenon called "failure to accept the task", first reported by Henle and Micheal (1956) and studied later by Richrer (1957): participants evaluated the content of the conclusion rather than the logical form of the argument when being asked to do a logical task. Richer (1957) suggested that this failure to accept the task might be due to "a general failure to grasp the concept of 'logical validity' " or one's "specific inability to differentiate 'logical validity' from other attributes of syllogisms" (p. 341). Consequently, participants were not performing "logical reasoning" as expected by the experimenter. In moral reasoning tasks, participants seem to make judgments based on their beliefs in the outcome possibilities instead of the information provided in the task instruction. Participants may experience difficulty in differentiating a hypothetical moral scenario from a real-world incident. They make their moral choices based on what they perceive as reasonable or consistent with their perceptions about the reality. This finding may highlight a weakness of moral reasoning studies in which participants are likely to feel the scenarios are unreal.

Finally, the investigation of expected values (calculated under the assumption that the probability judgments were not distorted) revealed that many participants chose $\tilde{K}$ even when their expected outcome values of $\tilde{K}$ was smaller than the ones of K. This suggests that factors other than outcome probabilities could affect moral choices in different dilemmas. One possible explanation is that participants' perceived "protected values" differ between the two dilemmas (Baron & Spranca, 1997). "Protected values" refer to the values that are against the trade-off in other type (economic/outcome) of values between the two choices. Taking a choice (usually an action such as killing) is at the cost of the "protected value" in addition to the death of the individual(s) given that choice. Participants might endorse higher protected values against killing for the hostage dilemma than they did for the car dilemma. Nevertheless, more research is needed to examine how moral decisions can be the joint product of both perceived outcome probabilities and protected values.

Overall, the present studies indicated that choosing $\tilde{K}$ does not necessarily entail that people engage in deontological reasoning, and choosing K does not necessarily entail that people engage in consequentialist reasoning. Preferring a choice (e.g., K) in dilemma A (e.g., switch) more than dilemma B (e.g., footbridge) could be because participants perceive the positive (or negative) outcome given that choice is more (or less) likely in A than in B. Participants may not accept the instructions in the descriptive moral reasoning task and may not conduct moral reasoning with the information provided as per experimenters' expectations. Without controlling for the equivalence of the outcome probabilities perceived by participants, it could be inadequate to derive conclusions such as that the discrepancy in moral choices between two dilemmas is because one dilemma induces higher emotional arousal than the other. Robustness of current findings, however, should be examined in future research with the inclusion of more variants of moral dilemmas.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press

References

Baron, J. (2013). Moral judgment: Acts, omissions, and rules. *Unpublished manuscript, Department of Psychology, University of Pennsylvania, Philadelphia, United States*.

Baron, J.,& Spranca,M. (1997). Protected values. *Organizational Behavior and Human Decision Processes, 70*, 1–16.

Crockett, M. J. (2013). Models of morality. *Trends in Cognitive Sciences*, *17*(8), 363–6. http://dx.doi.org/10.1016/j.tics.2013.06.005.

Cummins, D. D., & Cummins, R. C. (2012). Emotion and Deliberative Reasoning in Moral Judgment. *Frontiers in Psychology*. http://dx.doi.org/10.3389/fpsyg.2012.00328.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, *17*(12), 1082–1089. http://dx.doi.org/10.1111/j.1467-9280.2006.01834.x.

Greene, J. D., Cushman, F. a., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*(3), 364–371. http://dx.doi.org/10.1016/j.cognition.2009.02.001.

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*(3), 1144–1154. http://dx.doi.org/10.1016/j.cognition.2007.11.004.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*, 389–400. http://dx.doi.org/10.1016/j.neuron.2004.09.027.

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

Haidt, J. (2007). The new synthesis in moral psychology. *Science (New York, N.Y.)*, *316*(5827), 998–1002. http://dx.doi.org/10.1126/science.1137651.

Henle, M., & Michael, M. (1956). The influence of attitudes on syllogistic reasoning. *The Journal of Social Psychology, 44*(1), 115–127.

Kortenkamp, K. V., & Moore, C. F. (2014). Ethics under uncertainty: The morality and appropriateness of utilitarianism when outcomes are uncertain. *The American Journal of Psychology, 127*(3), 367-382.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. (2014). Emotion and Decision Making. *Annual Review of Psychology*. http://dx.doi.org/10.1146/annurev-psych-010213-115043.

Moll, J., & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive Sciences*, *11*, 319–321. http://dx.doi.org/10.1016/j.tics.2007.06.001.

Richter Jr, M. N. (1957). The theoretical interpretation of errors in syllogistic reasoning. *The Journal of Psychology, 43*(2), 341-344.

Rogers, J., Viding, E., & Chamorro-Premuzic, T. (2013). Instrumental and disinhibited financial risk taking: Personality and behavioural correlates. *Personality and Individual Differences*, *55*, 645–649. http://dx.doi.org/10.1016/j.paid.2013.05.011.

Shenhav, A., & Greene, J. D. (2014). Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *The Journal of Neuroscience[202F?]: The Official Journal of the Society for Neuroscience*, *34*(13), 4741–9. http://dx.doi.org/10.1523/JNEUROSCI.3390-13.2014.

Shou, Y., & Song, F. (2015). Moral reasoning as probability reasoning. In D. C. Noelle & P. P. Maglio (Eds.), *Proceedings of the 37th annual meeting of the cognitive science society* (pp. 2176–2181). Austin, TX.

Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, *17*(6), 476–7. http://dx.doi.org/10.1111/j.1467-9280.2006.01731.x.

Youssef, F. F., Dookeeram, K., Basdeo, V., Francis, E., Doman, M., Mamed, D., ... Legall, G. (2012). Stress alters personal moral decision making. *Psychoneuroendocrinology*, *37*(4), 491–8. http://dx.doi.org/10.1016/j.psyneuen.2011.07.017.

Appendix: ANOVA/F statistics for the contribution of each factor in Study 2.

Model	Scope		*F*	*p*
	DV: Negative Outcome Probability			
M0	Intercept Only			
M0a	M0 + Dilemma	M0 vs M0a	< 1	.912
M0b	M0 + Choice	M0 vs M0b	47.72	< .001
M0c	M0 + Condition	M0 vs M0c	3.02	.050
M1	M0 + Dilemma + Choice + Condition			
M1a	M1 + Dilemma×Choice	M1 vs M1a	8.75	.003
M1b	M1 + Dilemma× Condition	M1 vs M1a	2.43	.089
M1c	M1 + Choice× Condition	M1 vs M1a	<1	.595
M2	M1 + Dilemma + Choice + Condition + Dilemma× Condition + Dilemma× Condition + Choice× Condition			
M2a	M2 + Dilemma×Choice× Condition	M2 vs M2a	<1	.601
	DV: Positive Outcome Probability			
M0	Intercept Only			
M0a	M0 + Dilemma	M0 vs M0a	184.06	< .001
M0b	M0 + Choice	M0 vs M0b	6.82	.009
M0c	M0 + Condition	M0 vs M0c	<1	.922
M1	M0 + Dilemma + Choice + Condition			
M1a	M1 + Dilemma×Choice	M1 vs M1a	.97	.002
M1b	M1 + Dilemma× Condition	M1 vs M1a	<1	.723
M1c	M1 + Choice× Condition	M1 vs M1a	<1	.700
M2	M1 + Dilemma + Choice + Condition + Dilemma× Condition + Dilemma×Choice + Choice× Condition			
M2a	M2 + Dilemma×Choice× Condition	M2 vs M2a	<1	.997

Note. Choice is Given Choice.

The first half of the table shows the model fit comparison when the DV is the probability judgments for the negative outcomes for the two choices in the two dilemmas, while the DV in models of the second half of the table is for the positive outcomes.

The first column indicates the number of each model. The second column is the model scope, which indicates what terms were included in the model. For example, M0 only contains the intercept (M0: DV˜ Intercept); M0a had an additional term "Dilemma" comparing to M0 (M0a: M0 + Dilemma; DV˜ Intercept + Dilemma); M1 had three main effect terms ("Dilemma", "Choice", and "Condition") comparing to M0 (M1: M0 + Dilemma + Choice + Condition; DV˜ Intercept + Dilemma + Choice + Condition). The next column indicates which two models are compared for the F test. For example, "M0 vs M0a" means the *F* test is based on a comparison of M0a (DV˜ Intercept + Dilemma) and M0 (DV˜ Intercept). A significant F statistic for "M0 vs M0a" indicates the term "Dilemma" had a significant contribution to the model fit of a model without this term.

https://doi.org/10.1017/S1930297500006501 Published online by Cambridge University Press