

Sacrificing Oneself or Another: The Difference Between Prescriptive and Normative Judgments in Moral Evaluation

Psychological Reports
0(0) 1–23

© The Author(s) 2020
Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0033294119896061

journals.sagepub.com/home/prx



Štěpán Bahník 

The Prague College of Psychosocial Studies, Prague,
Czech Republic

Emir Efendic

Catholic University of Louvain, Psychological Sciences
Research Institute, Louvain-La-Neuve, Belgium

Marek A. Vranka

Faculty of Social Sciences, Charles University, Prague,
Czech Republic

Abstract

When asked whether to sacrifice oneself or another person to save others, one might think that people would consider sacrificing themselves rather than someone else as the right and appropriate course of action—thus showing an other-serving bias. So far however, most studies found instances of a self-serving bias—people say they would rather sacrifice others. In three experiments using trolley-like dilemmas, we tested whether an other-serving bias might appear as a function of judgment type. That is, participants were asked to make a prescriptive judgment (whether the described action should or should not be done) or a normative judgment (whether the action is right or wrong). We found that participants exhibited an other-serving bias only when asked whether self- or other-sacrifice is *wrong*. That is, when the

Corresponding Author:

Štěpán Bahník, The Prague College of Psychosocial Studies, Hekrova 805, Prague 149 00, Czech Republic.
Email: bahniks@seznam.cz

judgment was normative and in a negative frame (in contrast to the positive frame asking whether the sacrifice is *right*). Otherwise, participants tended to exhibit a self-serving bias; that is, they approved sacrificing others more. The results underscore the importance of question wording and suggest that some effects on moral judgment might depend on the type of judgment.

Keywords

Moral judgment, self-other asymmetry, trolley dilemma, self-sacrifice

Introduction

When making judgments, people exhibit various asymmetries in how they evaluate themselves compared to how they evaluate others. Usually, these evaluations are self-serving. For instance, people see positive attributes as more descriptive of themselves than of others (Brown, 1986), they expect others to be more affected by alcohol than themselves (Rohsenow, 1983), and they see many more operational, as well as cognitive, biases in others than in themselves (Pronin, Lin, & Ross, 2002). People also tend to believe that they are, compared to others, less likely to experience negative and more likely to experience positive events (Weinstein, 1987).

Similar asymmetries have also been observed in moral judgments. For example, in the case of moral hypocrisy, people judge moral transgressions committed by themselves less harshly than the same acts committed by others (Valdesolo & DeSteno, 2007, 2008), and they consider themselves to be generally more ethical, charitable, cooperative, and fair than others (Epley & Dunning, 2000).

However, asymmetries in moral judgments do not always have to be self-serving. In some cases, people might exhibit other-serving biases—their judgments can be biased in favor of others, rather than themselves. For example, Crockett, Kurth-Nelson, Siegel, Dayan, and Dolan (2014) found that when participants could pay to reduce pain by electrical shocks, they were willing to pay more to reduce others' pain than their own, suggesting that people value their own suffering less than the suffering of others.

Similarly, one could expect that in studies using trolley (and similar) dilemmas, some people might consider self-sacrifice as the morally right action. These kinds of dilemmas, named after the most famous example by Foot (1967), present a scenario in which a larger number of people can be saved by sacrificing another person. In the trolley dilemma, a runaway trolley can be shifted from a track with five workers to a track with just one worker. Within the landscape of moral theories that purport to assess our ethical choices, two perspectives are most commonly evoked: the deontological and the consequential.

The deontological perspective speaks on which choices are morally required or permitted and is akin to a rule-based system. From the deontological perspective, it is wrong to use someone as a means to an end, ignoring their rights—e.g., sacrifice the worker in the trolley dilemma—even when doing so prevents others dying. Letting the trolley run over the five workers is therefore considered the morally right action from the deontological perspective. However, the deontological position differs when instead of an anonymous worker one sacrifices oneself—even though self-sacrificing in an attempt to save the lives of others is not morally obligatory, it would not be usually considered morally wrong and it can be even seen as praiseworthy (Jacobs, 1987; Slote, 1984).¹ Consequentialism, in contrast, holds that moral choices ought to be assessed purely on the end states that they bring about. Sacrificing one worker to save five brings about more good overall (i.e., saves more lives), and it is therefore considered the morally right action from the perspective of consequentialism. By extension, from this point of view, self- and other-sacrifice should be evaluated equivalently, as in the end the number of lost and saved lives is the same. However, some philosophers argue that the overall amount of harm or good done is not the only thing that matters when such actions are evaluated (Slote, 1984; Stocker, 1976). Discounting one's own pleasures and pains, and thus exhibiting an other-serving bias, seems to be morally permissible and even praiseworthy (Splawn, 2001). For example, although the action of stabbing causes harm and can therefore be considered bad, it seems difficult to judge stabbing oneself and stabbing someone else as being morally equivalent. In line with this reasoning, some research findings (Lester, 2014; Sachdeva, Iliev, Ekhtiari, & Dehghani, 2015) suggest that some people do in fact think that self-sacrifice is the more appropriate action in the trolley dilemma—thus exhibiting an other-serving bias.

How people actually evaluate self- and other-sacrifice in the trolley dilemma was investigated by Huebner and Hauser (2011). In their modified version of the trolley scenario, participants had to decide which of the following three actions they should do: turn a trolley to a track with an anonymous person, let the trolley run over five people by doing nothing, or turn the trolley to a track where they were standing themselves. While one third of participants answered that they should sacrifice themselves, displaying an other-serving bias, almost half of the participants said that they should save the five people by turning the trolley to the anonymous person, showing a self-serving bias.

We suspect that the type of judgment people were asked to make may help to explain why people were more likely to display a self-serving bias. As shown in previous studies, question wording may substantially influence moral judgments. Psychological studies have distinguished between so-called prescriptive and normative types of judgments. A prescriptive moral judgment relates to what people *should do* while a normative judgment relates to what is *right* or *wrong* to do. For example, Barbosa and Jiménez-Leal (2017) found that people were more condemning of the sacrifice in the trolley dilemma when asked for a

prescriptive judgment than when asked for a normative judgment. More standard philosophical distinctions have been made between normative and descriptive judgments (i.e., judgments about what people believe they or someone else *would* actually do; Tassy, Oullier, Mancini, & Wicker, 2013). Related to this distinction, studies showed lower approval of sacrifice in the trolley dilemma in case of normative, rather than descriptive judgments, and lower approval of aggressive retaliation after being provoked in case of prescriptive, rather than descriptive judgments (Goldstein, Tisak, & Boxer, 2002).

In the Huebner and Hauser's (2011) study, participants made prescriptive moral judgments rather than normative judgments, which are usually made by philosophers and lay persons arguing morality of self-sacrifice. These two types of judgments differ in significant ways – while normative judgments are value judgments and are likely to be based solely on the morality of the action, prescriptive judgments require additional inference from values to a recommendation of an action. Prescriptive judgments may therefore include consideration of other factors such as reputational consequences, one's self-interest, other norms, or the probability of achieving the intended outcome (Elqayam, Thompson, Wilkinson, Evans, & Over, 2015; Elqayam, Wilkinson, Thompson, Over, & Evans, 2017).

Therefore, we believe that the lack of an other-serving bias in Huebner and Hauser (2011) can be explained by the type of judgment the participants were asked to make. Specifically, we hypothesize that people will exhibit an other-serving bias when the judgment is normative and therefore more likely to be solely based on the morality of the action in question.² On the other hand, when making a prescriptive judgment, people should be more likely to take their self-interest into account and thus exhibit a self-serving bias. The following studies aim not only to reconcile the currently conflicting empirical results and philosophical intuitions about moral judgments of self-sacrifice, but also to shed more light on how different question prompts may influence moral judgments in general.

Study 1a

Methods

Participants. A total of 198 German-speaking participants (72% female, median age 23) from the University of Würzburg subject pool took part in the study. The study was part of several unrelated studies administered during a single session. Participants were paid €5 for participation in the whole session, which lasted about 40 minutes.

Materials

We adapted four trolley-like dilemmas from Moore, Clark, and Kane (2008). The dilemmas presented situations where participants had to decide whether to

sacrifice one person to save more people. In all four dilemmas, the sacrifice of the person was a side effect of the action which saved the other people. The action that led to the death of the sacrificed person was also always impersonal, such as releasing a hatch or pushing a button; that is, there was no direct contact with the sacrificed person (Greene et al., 2009). In these aspects, the dilemmas shared the structure of the commonly used basic trolley dilemma. Each dilemma had 2×2 versions differing in whether participants sacrificed themselves or another person (Target of sacrifice: Self vs. Other) and whether they judged if the described action was wrong or if they should perform it given the circumstances (Type of judgment: Wrong—normative vs. Should—prescriptive). While the person to be sacrificed differed based on the scenario, the actor making the decision about the sacrifice in the scenario was always the participant. For example, the submarine dilemma read as follows (changes in different versions are in brackets):

You are a crewmember on a marine-research submarine traveling underneath a large iceberg. An onboard explosion has damaged the boat, killed and injured several crewmembers. Additionally, it has caused a collapse of the only access corridor between upper and lower parts of the boat. The upper section, where most of the crew is located, does not have enough remaining oxygen for all to survive until the submarine reaches the surface. Only you are located in the lower section, where there is enough oxygen. [You and one other crewmember are located in the lower section, where there is enough oxygen.]

There is an emergency access hatch between the upper and lower sections of the ship. If released by an emergency switch, to which only you have access, it will fall to the deck and allow oxygen to reach the upper section of the submarine.

However, the hatch will crush you, since you are trapped and lying beneath it. [However, the hatch will crush the crewmember in the lower section, since he is trapped and lying beneath it.]

All of the crew in the upper section of the submarine are almost out of air though, and all will die if you do not do this. On the other hand, if you do nothing, you will survive and will be rescued on the other side of the iceberg. [On the other hand, if you do nothing, the crewmember will survive and you will be rescued on the other side of the iceberg.]

Should you release the hatch? (*prescriptive*)

[Is it wrong to release the hatch? (*normative*)]

The remaining three dilemmas, alongside their German wordings used in the experiment, are presented at: <https://osf.io/n7e8c/wiki/home/>. All materials

were prepared based on the original English dilemmas (Moore et al., 2008) by a native German speaker fluent in English and checked by another German speaker.

Procedure and design

The experiment was administered in a laboratory setting. Participants did the study on computers in groups of up to 8. They were separated by barriers and could not interact. Participants first read short instructions that explained the task and then made four judgments corresponding to the four dilemmas which were presented separately and in a fixed order. We used the “counterbalanced design” (Westfall, Kenny, & Judd, 2014) where each participant was presented with all four dilemmas, but each dilemma was in a different version, corresponding to the four possible conditions (Target of sacrifice: Self vs. Other \times Type of judgment: Wrong vs. Should). Each participant was presented with each possible combination exactly once (in a random order) meaning that it was not possible for participants to see two dilemmas in the same version. Given that order effects were a possible issue, this design allowed us to analyze the first scenario, where no order effect could be present, separately with sufficient power. Participants answered whether they considered the described action wrong or whether they thought that they should do the action by clicking on a Yes or No button. The binary answer served as the dependent variable. The entire experiment was conducted in German.

Results

Data analysis was done using multilevel logistic regression (Gelman & Hill, 2006) with the data for the normative judgment reverse coded—positive values thus indicated approval of a sacrifice of a person for the sake of a higher number of people. Dilemmas and participants were treated as random factors, allowing for generalization of results to other similarly sampled dilemmas as well as other participants (Judd, Westfall, & Kenny, 2012). Effects of the type of judgment, the target of sacrifice, and their interaction were included as predictors. Apart from random intercepts for dilemmas and participants, the model included random slopes for dilemmas.

Results showed that there was a general main effect of the target of sacrifice. Participants were less likely to approve of sacrificing someone else than themselves, $z = -2.90$, $p = .004$, odds ratio (OR) = 0.30, 95% confidence interval (CI) [0.13, 0.68]. There was also a main effect of type of judgment, in that participants were less likely to approve of the sacrifice in the prescriptive than in the normative condition, $z = -5.43$, $p < .001$, OR = 0.25, 95% CI [0.15, 0.41]. Most importantly, the interaction between the two factors was also significant, $z = 3.00$, $p = .003$, ratio of OR = 3.53, 95% CI [1.55, 8.05], showing that the

self-other asymmetry was more pronounced in the case of normative than in the case of prescriptive judgment. This can be seen when the analysis was done separately for the two types of judgments. Whereas for the “wrong” question, the sacrificed person effect was significant, $z = -2.92, p = .004, OR = 0.32, 95\% CI [0.15, 0.69]$, it was not significant for the “should” question, $z = 0.26, p = .80, OR = 1.06, 95\% CI [0.69, 1.63]$.

When the analysis was conducted with separate logistic regressions for each dilemma, the interaction between the target of sacrifice and the type of judgment was significant for the last two dilemmas, but not significant for the first two dilemmas.³ Separate results for the four dilemmas are presented in Figure 1.

The figure displays proportions of participants indicating their approval of sacrificing a person to save more people (i.e., the proportion of people answering “yes” to the “should” question and “no” to the “wrong” question). The order by rows corresponds to the order of presentation of the dilemmas. The error bars represent 95% CIs.

Discussion

The first study showed that there is a difference between normative judgments of wrongness for self- and other-sacrifice. People were more likely to say that it is

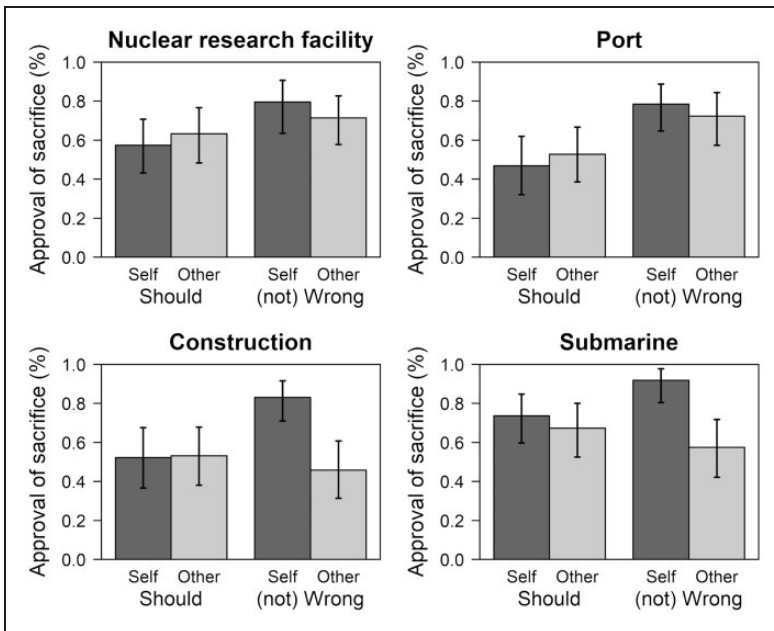


Figure 1. Approval of sacrifice by dilemma in Study 1a.

wrong to sacrifice someone else than themselves. On the other hand, no difference was found in prescriptive judgments of what should be done. The results support our hypothesis that prescriptive and normative moral judgments do not necessarily correspond to each other. However, unlike Huebner and Hauser (2011), we did not observe a self-serving bias in the prescriptive condition.

To replicate the results, we conducted a follow-up study with a different sample of participants, which also addressed a possible issue with the order of presentation of the dilemmas. Specifically, we found a significant interaction between the target of sacrifice and the type of judgment only for the last two dilemmas in Study 1a, so we reversed the order of dilemmas to examine whether the results could be influenced by their order of presentation (cf. Di Nucci, 2013).

Study 1b

Methods

Participants. A total of 357 US MTurk workers participated in the study. Participants were paid 30¢ for participation, which lasted about 4 minutes. We did not ask the participants for their demographic information; however, MTurk workers are generally more diverse than student samples, although still not entirely representative of the US population. They tend to be younger and more educated than the general US population (Paolacci & Chandler, 2014).

Materials and procedure

We used the same materials and design as in Study 1a. The only difference being that the dilemmas were presented in their original English form (Moore et al., 2008). The experiment was administered online using a custom written web application. Since the effect seemed to be stronger for the last two dilemmas in Study 1a, we reversed the order of the dilemmas. Thus, the last two dilemmas from Study 1a were now the first two dilemmas in this study. We kept the order of scenarios constant for all participants in order to have sufficient power to analyze the data from the scenario presented first, which was unaffected by other scenarios.

Results

Based on a pre-registered exclusion criterion, 82 participants were removed from the analysis because they answered at least one question in less than 15 seconds. Two further participants were excluded because they did not answer all the questions, leaving us with 273 participants.

The same analysis procedure as in Study 1a was used. The results again showed a main effect of the target of sacrifice in that people were less likely

to approve of sacrificing someone else than themselves, $z = -2.73$, $p = .006$, $OR = 0.49$, 95% CI [0.30, 0.82], and a main effect of the type of judgment in that they were less likely to approve of sacrificing the person in the prescriptive than in the normative condition, $z = -3.56$, $p < .001$, $OR = 0.39$, 95% CI [0.23, 0.65]. Furthermore, the interaction between the two factors was again significant, $z = 2.06$, $p = .04$, ratio of $OR = 2.07$, 95% CI [1.04, 4.14], providing additional evidence that the self-other asymmetry is more pronounced in normative than in prescriptive judgment.

Separate analyses for the two types of judgment again yielded the same results. While there was no effect of the target of the sacrifice when people were asked what should be done, $z = 0.15$, $p = .88$, $OR = 1.03$, 95% CI [0.70, 1.50], people were less likely to say that it is wrong to sacrifice themselves than someone else, $z = -3.03$, $p = .002$, $OR = 0.51$, 95% CI [0.33, 0.79].

When the four scenarios were analyzed separately, a significant interaction was found for the first two scenarios (the same two as in the first study).⁴ Separate results for the four dilemmas are presented in Figure 2.

The figure displays proportions of participants indicating their willingness to sacrifice a person to save more people (i.e., the proportion of people answering “yes” to the “should” question and “no” to the “wrong” question). The order by

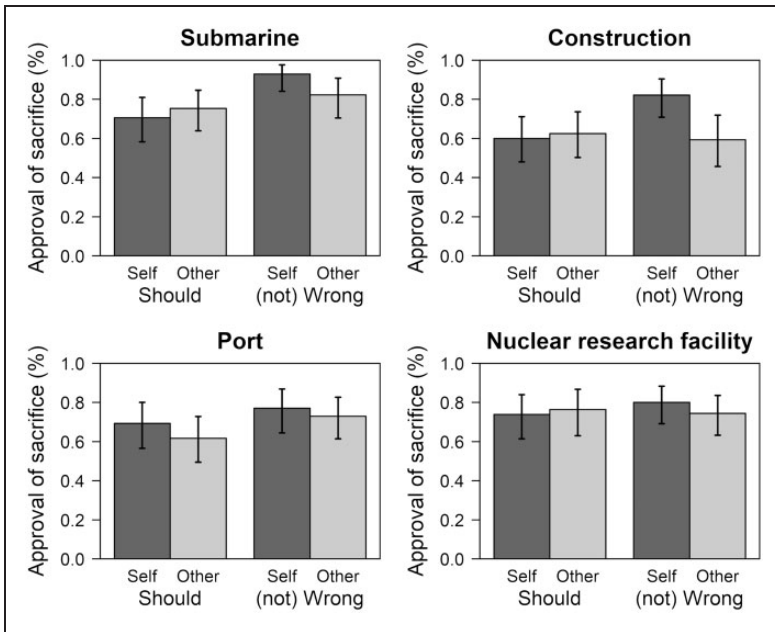


Figure 2. Approval of sacrifice by dilemma in Study 1b.

rows corresponds to the order of presentation of the dilemmas. The error bars represent 95% CIs.

Discussion

Study 1b replicated the results of Study 1a with a different sample of participants and language of materials. We again found the predicted self-other asymmetry. Specifically, while there was no difference in prescriptive judgment of whether one should sacrifice oneself or someone else, participants showed a difference in normative judgment—they were more likely to say that it is wrong to sacrifice someone else than themselves. Taken together, the two studies showed no evidence of an order effect, which is in line with recent findings showing that order effects in moral judgment may be relatively uncommon and limited to a narrow set of circumstances (Wiegmann, Okan, & Nagel, 2012). However, the results suggest that the strength of the other-serving bias may depend on used scenarios since we found evidence for the effect only in two dilemmas. Finally, unlike in Huebner and Hauser (2011) and again as in Study 1a, we did not observe a self-serving bias in the prescriptive condition.

However, one could argue that the obtained results are not caused by different types of judgments, but that they are due to the specific wording that was used. Namely, both studies used negative framing (“is it wrong?”) for the normative judgment, and positive framing (“should you do it?”) for the prescriptive judgment. That is, an affirmative answer meant disapproval of sacrifice in the normative judgment and approval of sacrifice in the prescriptive judgment. The difference between positive and negative framing and not the difference between the two types of judgment might have driven the results. In the next study, we therefore explicitly manipulated the framing in order to address this issue. Alongside the original questions of whether the action is wrong and whether it should be done, we used a positively framed normative judgment (“is it right?”) and a negatively framed prescriptive judgment (“should it not be done?”).

Additionally, we explored whether the results of the first two studies replicate even when the dilemmas are written in the third-person perspective. Therefore, we added a perspective factor, by which we varied whether it was the participants themselves who were deciding whether to sacrifice themselves or someone else, or if it was another person named “John.” Previous studies have demonstrated effects of such manipulations of perspective on moral judgment. For example, Huebner and Hauser (2011) found that slightly fewer people reported that they would sacrifice themselves when their trilemma scenario was presented in the first-person perspective than in the third-person perspective. Using trolley dilemmas more similar to ours, Sachdeva et al. (2015) found an other-serving bias when the scenario was presented in the first-person perspective but not when it was presented in the third-person perspective.

Actions and their causes tend to be perceived differently based on whether people perform them themselves or whether they only observe them (Malle, 2006). Along these lines, Nadelhoffer and Feltz (2008) found that people were more likely to judge as morally permissible to not sacrifice the person in the trolley dilemma when the scenario was written from the third-person perspective than when it was written from the first-person perspective—possibly because people experience stronger negative emotions when considering sacrifice while imagining being an agent rather than an observer (Miller, Hannikainen, & Cushman, 2014). However, these negative feelings might be absent when one is considering self-sacrifice, in which no other person is harmed. Based on this reasoning, we expected that the other-serving bias from the first two studies would be more pronounced in dilemmas presented from the first-person perspective.

Study 2

Methods

Participants. A total of 1892 participants filled a pen-and-paper questionnaire (which asked participants mainly about university and high school education) before the General Academic Prerequisites (GAP) test. The GAP test is used for university admission at the undergraduate level by a wide range of universities in the Czech Republic. Participants were informed that the questionnaire was fully voluntary, was going to be analyzed anonymously, and could not affect the results of the GAP test in any way. The scenario used in the present study was included in the questionnaire. We excluded 314 participants for failing to respond to the question, leaving us with 1578 participants. In the remaining sample, 54.7% of participants were women and the median age was 19.4 years (interquartile range = 1.49).

Materials and design

Only the submarine dilemma scenario from the previous studies was used. The dilemma was translated to Czech by a native Czech speaker fluent in English, and the wording was slightly adjusted so that it made grammatical sense in all the conditions. Furthermore, the questions were reworded to statements (see full materials at <https://osf.io/n7e8c/wiki/home/>). All materials were checked by another researcher fluent in both Czech and English. We asked participants about their agreement with the statements. The responses were provided on a six-point Likert scale ranging from 1 (completely agree) to 6 (completely disagree). We substituted the binary response dependent variable from the first study for a continuous scale in order to obtain more fine-grained responses. Responses for the positively framed scenarios were reversed such that higher

values always corresponded to higher approval of the sacrifice (similarly as in Study 1).

We used a between-subject design in Study 2. There were 16 versions of the dilemma, resulting from a combination of four factors. In addition to the two factors that were the same as in Study 1 (i.e., Type of judgment: Normative vs. Prescriptive, and Target of sacrifice: Self vs. Other), we added the factor of framing (Positive vs. Negative) and the factor of perspective (“You” vs. “John”). The framing factor manipulated whether the statement participants were asked to judge was framed such that agreement with it meant approval (“should,” “it is right”) or disapproval of the sacrifice (“should not,” “it is wrong”). Unlike in Study 1, the wording of the normative statement also included a reference to the actor (e.g., “It is wrong *for me* to release the hatch”). The perspective factor manipulated who is the actor in the scenario (“you” vs. “John”). The scenario was thus presented in the second or third person and the evaluated statement in the first or third person (e.g., “I should” vs. “John should”). The participants were randomly assigned to a single version of the scenario.⁵

Results

Full model. All four factors as well as all their interactions were inserted into a regression model.⁶ All the predictor variables were mean centered. The full results are presented in Table 1 (see also Figure 3).⁷ They show that people approved of the sacrifice more in the positive frame (mean (M) = 4.16, standard deviation (SD) = 1.18) than in the negative frame (M = 3.68, SD = 1.31), and that participants were more willing to sacrifice someone else (M = 4.01, SD = 1.21) than themselves (M = 3.84, SD = 1.31), suggesting a general self-serving bias. Most importantly, the model revealed a significant three-way interaction between the framing, target of sacrifice, and type of judgment so we decomposed the interaction by frames. Since there was no significant effect of the perspective factor, the results are pooled across both perspectives.

The figure displays the average rating of the approval of sacrificing one person to save more people (i.e., reverse scored disagreement with positively framed questions). The error bars represent 95% CIs. Individual data with applied jitter are shown as transparent points.

Positive frame. We first took a closer look at the positive frame condition. Similar to Study 1, we focused on how the effect of type of judgment (Prescriptive vs. Normative) might differ based on the target of judgment (Self vs. Other). The results indicated that, in the positive frame, participants approved more of sacrificing the other person (M = 4.27, SD = 1.11) than themselves (M = 4.04, SD = 1.24), suggesting a self-serving bias, irrespective of whether the judgment was described as prescriptive or normative (see Table 1).

Table 1. Regression coefficients and 95% confidence intervals for the full model and two models separated by the frame factor for Study 2.

| | Full model (N = 1578) | | Positive frame (N = 809) | | Negative frame (N = 769) | |
|--|-----------------------|----------------|--------------------------|----------------|--------------------------|---------------|
| | b | 95% CI | b | 95% CI | b | 95% CI |
| Intercept | 3.93** | [3.87, 3.99] | 4.16** | [4.08, 4.24] | 3.68** | [3.60, 3.77] |
| Framing | 0.48** | [0.35, 0.60] | | | | |
| Judgment Type | 0.08 | [-0.04, 0.21] | 0.09 | [-0.07, 0.25] | 0.07 | [-0.11, 0.26] |
| Perspective | -0.04 | [-0.17, 0.08] | | | | |
| Sacrifice Target | -0.18** | [-0.30, -0.06] | -0.24** | [-0.40, -0.08] | -0.11 | [-0.30, 0.07] |
| Framing × Judgment Type | 0.01 | [-0.23, 0.26] | | | | |
| Framing × Perspective | -0.12 | [-0.37, 0.12] | | | | |
| Judgment Type × Perspective | -0.17 | [-0.41, 0.08] | | | | |
| Framing × Sacrifice Target | -0.13 | [-0.37, 0.12] | | | | |
| Judgment Type × Sacrifice Target | 0.11 | [-0.14, 0.35] | -0.19 | [-0.51, 0.14] | 0.43* | [0.06, 0.80] |
| Perspective × Sacrifice Target | 0.09 | [-0.15, 0.34] | | | | |
| Framing × Judgment Type × Perspective | 0.01 | [-0.48, 0.50] | | | | |
| Framing × Judgment Type × Sacrifice Target | -0.62* | [-1.11, -0.13] | | | | |
| Framing × Perspective × Sacrifice Target | -0.11 | [-0.60, 0.38] | | | | |
| Judgment Type × Perspective × Sacrifice Target | 0.34 | [-0.15, 0.83] | | | | |
| Framing × Judgment Type × Perspective × Sacrifice Target | -0.40 | [-1.38, 0.58] | | | | |

Note: The models decomposed by framing do not include perspective or its interactions as predictors given that they had no effect in the full model. CI: confidence interval.

* $p < .05$; ** $p < .01$.

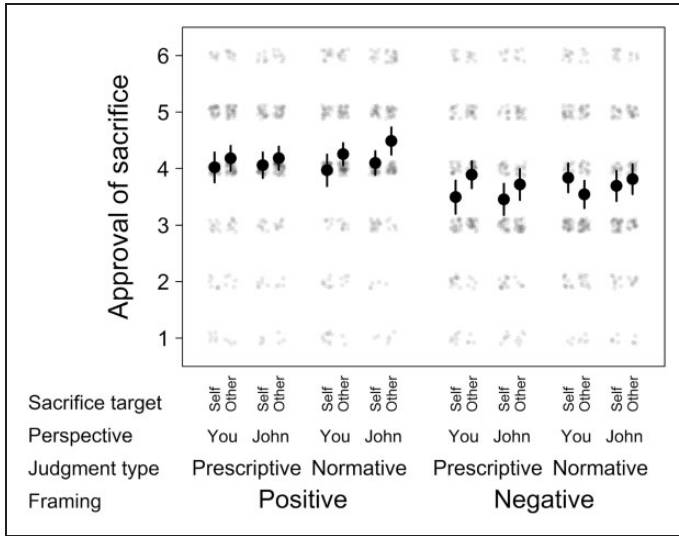


Figure 3. Approval of sacrifice by condition in Study 2.

Negative frame. We then took a closer look at the negative frame condition. Similar to above, we focused on how the effect of type of judgment (Prescriptive vs. Normative) might differ based on the target of sacrifice (Self vs. Other). The results indicated an interaction between the type of judgment and the target of sacrifice (see Table 1). Separate analyses for the two types of judgments show that, when the judgment was prescriptive, participants agreed more that they should not sacrifice themselves ($M = 3.48$, $SD = 1.37$) than the other person ($M = 3.81$, $SD = 1.25$), $b = -0.34$, $p = .014$, 95% CI $[-0.60, -0.07]$, suggesting a self-serving bias. When the judgment was normative, on the other hand, participants agreed more that it is wrong to sacrifice someone else ($M = 3.67$, $SD = 1.26$) than themselves ($M = 3.80$, $SD = 1.33$). However, this difference, although in the same direction as in Study 1, was not significant, $b = 0.10$, $p = .45$, 95% CI $[-0.16, 0.35]$. Even though the interaction between the target of sacrifice and perspective was not significant, we next analyzed the effect of target of sacrifice only for the first-person perspective for the normative judgment in the negative frame. Because this condition corresponds directly to the materials used in Study 1, this amounted to a direct replication of the normative judgment condition from Study 1. The results show that participants agreed more that it is wrong to sacrifice someone else ($M = 3.54$, $SD = 1.26$) than themselves ($M = 3.84$, $SD = 1.32$), but the difference was not significant, $b = 0.30$, $p = .10$, 95% CI $[-0.06, 0.65]$ (see Figure 3).

Discussion

As in Study 1, we again found evidence of self-other asymmetry, although not in accord with all our predictions. Overall, participants showed a general self-serving bias. More interestingly though, framing of the questions had a significant impact on the results. Separate analyses of the two frames showed that when the questions were positively framed, participants displayed a self-serving bias in their judgments. That is, they were more likely to approve of sacrificing someone else than themselves, independent of the type of judgment. When the questions were negatively framed, participants were more likely to exhibit a self-serving bias when the judgment was prescriptive. On the other hand, when the judgment was normative, there was a similar tendency for an other-serving bias as in Study 1. However, the effect was not significant—possibly due to lower statistical power associated with a between-subject design or the use of a different response scale. Some variability of an observed effect size can also be expected even when the effect is true. These results support and extend the findings of Study 1.

The fact that we did not observe a self-serving bias in Study 1 might have been caused by using only a positively framed prescriptive condition in which the bias is less pronounced even in the current study (see Figure 3). While the two types of judgments produced different effects, it must be highlighted that the differences seemed to be qualified by framing. Framing and judgment type seem to be interdependent and future investigations of moral judgments should take this into account.

Surprisingly, there was no significant effect of perspective on moral evaluations—whether it was the participants themselves or someone else (i.e., “John”) who was making the decision in the scenario did not seem to make a difference. Our specific prediction was that the other-serving bias would be more pronounced in the first-person condition. Similarly as in the study by Sachdeva et al. (2015), the tendency for an other-serving bias was observed only in the first-person perspective. However, the interaction with the perspective factor was not significant, again possibly because of low statistical power.

General discussion

We are all familiar with the trope of a hero sacrificing himself or herself for the good of the many. Such an action would be an example of an other-serving bias, and it could be considered as supererogatory, that is, morally praiseworthy although not required (Chisholm, 1963). Despite the ubiquity of such actions in literature, previous psychological studies that looked at moral judgments mostly focused on, and have quite often found, examples of self-serving biases (Huebner & Hauser, 2011).

To explore this discrepancy, in this paper, we took a closer look at one factor that might be particularly important for studying self-other asymmetries in moral judgments—namely, whether the judgment in question was prescriptive (i.e., whether something should or should not be done) or normative (i.e., whether something is right or wrong). In three experiments, we found that people were more likely to approve of sacrificing themselves than others, but the last study showed that this was the case only when the judgment was framed negatively and in a normative way; that is, when participants were asked whether it is wrong to sacrifice themselves or someone else. Our results thus show an example of an other-serving bias in moral judgments and provide empirical support for intuitions described by Slote (1984) and Stocker (1976). However, it is worth highlighting that the bias in this direction seems to appear only in a very narrow set of conditions (i.e., only when judging whether the sacrifice is wrong) and is probably less common than would be expected. This is in line with recent findings of low willingness to undergo harm for the benefit of others in laboratory settings (Volz, Welborn, Gobel, Gazzaniga, & Grafton, 2017). On the other hand, Sachdeva et al. (2015) found a similar other-serving bias using a question asking whether sacrificing a person was appropriate. It is possible that the question about the appropriateness of a behavior is interpreted similarly to question about its wrongness, even if the frame is different in each case. Apart from the other-serving bias, we also observed the much more common self-serving bias in Study 2 (e.g., Ditto, Pizarro, & Tannenbaum, 2009; Huebner & Hauser, 2011; Valdesolo & DeSteno, 2007, 2008), which seemed to be present especially when participants judged whether the sacrifice is right and whether it should not be done.

The finding of an other-serving bias only in negative normative judgment could be explained by people being more focused on moral rules and prohibitions when deciding whether an action is wrong. Their judgment would therefore be less concerned with evaluation of specific consequences. In the present case, people might focus mainly on harm done to others, because they see normative moral rules as applying mostly to harm done to someone else (Gray & Wegner, 2011). Therefore, they would not view self-sacrifice as a violation of a moral rule and therefore judge it as not wrong. On the other hand, when thinking about what should and should not be done, these rules could be less salient and thus other factors such as consequences and self-interest could more likely affect the judgment.

The proposed explanation is also in accord with the differences demonstrated between descriptive and normative judgment by Tassy et al. (2013). The study found that people are more willing to approve of sacrificing a person to save more people when making a descriptive judgment (“would you?”) than normative judgment (“is it acceptable?”). Additionally, Tassy et al. found that affective proximity to the sacrificed person had a larger effect on descriptive judgment. It is possible that people are less influenced by affective proximity in normative

judgment because the moral rule not to kill is the same for strangers as well as for close relatives. However, the consequences of killing one's relative or a stranger are different. And, because the consequences are possibly more salient when judging what one would do (similarly as when judging what should be done), the difference in consequences influences the descriptive judgment more. Similarly, people making prescriptive judgments in our studies might have taken into consideration not only rules forbidding killing of others, but their own self-interest as well.

However, this explanation does not account for the unexpected lack of an other-serving bias in the positively framed normative judgment, that is, when participants judged whether the sacrifice is right. Although we originally expected the effect to be present for both judgments of wrong and right, differences between moral judgments in positive and negative frame have been observed before (e.g., Janoff-Bulman, Sheikh, & Hepp, 2009). One possible explanation for why we did not observe an other-serving bias in positively framed normative judgment could be the lack of positive deontological rules similar to those prohibiting harming others—as mentioned by Chisholm (1963), suffering for the benefit of others is considered praiseworthy, but not obligatory. In a case of positively framed normative judgment, the deontological intuition, responsible for the other-serving bias observed in the negative frame, is therefore absent, allowing other considerations such as self-interest to affect the judgment.

Apart from the interaction of framing with type of judgment and the target of sacrifice, we also found that people were generally more likely to approve the sacrifice when the question was asked in positive framing than when it was asked in negative framing. A possible explanation for this effect might be that a positively framed question is answered affirmatively when the sacrifice is considered obligatory as well as when it is deemed supererogatory (i.e., good, but not obligatory) and a negatively framed question is answered in the negative only when the sacrifice is considered obligatory (Chisholm, 1963). That is, people may agree that the sacrifice is the right thing to do, but would disagree that it is wrong not to do it when they consider the sacrifice supererogatory. This would lead to the main effect of framing we observed in the data. Another explanation could be that people are more likely to focus on reasons in favor of the sacrifice if the framing is positive but more likely to focus on reasons against the sacrifice when the framing is negative.

The observed effects of judgment type and framing underscore how question wording may influence moral judgment (Barbosa & Jiménez-Leal, 2017; O'Hara, Sinnott-Armstrong, & Sinnott-Armstrong, 2010; Pastötter, Gleixner, Neuhauser, & Bäuml, 2013; Tassy et al., 2013). It is possible that similar wording effects might limit generalizability of other previously found effects on moral judgment, and researchers should pay attention to wording effects in future studies to overcome this limitation.

Interestingly, we did not find an effect of perspective on judgments (i.e., whether the actor in the scenario was the participant or someone else). Nadelhoffer and Feltz (2008) demonstrated a type of “actor-observer” bias where people hold other people to different moral standards than themselves. They argue that the explanation of this effect relies on different affective responses people have in scenarios where they are the actor and where they are the observer. It could be that our scenarios failed to evoke these affective reactions. Consistently, Sachdeva et al. (2015), who found the effect of perspective on self-other asymmetry, used the footbridge dilemma, which is similar to the trolley dilemma, but involves direct contact with the sacrificed person (pushing a person onto the tracks), thus being more likely to evoke a strong affective response. However, our failure to obtain the effect of perspective is not an isolated one. Recently, Cova et al. (2018) and Andow (2018) have failed to replicate this effect in high powered direct and conceptual replication attempts. While there might be some subtle differences between our and the other studies, the findings that we obtained here in combination with the recent failed replication attempts invite caution in interpreting the effect of perspective on moral judgments.

The results appear to be quite robust as we obtained similar findings across laboratory, online, and pen-and-pencil settings. Furthermore, we collected data across three samples (German, US, and Czech) from different cultural milieus, each time using a different language. Despite these differences, the results did not significantly fluctuate across the samples. This might be due to the fact that the translation of the words crucial to our findings (i.e., “right,” “wrong,” “should”) was straightforward—the words have clear and delimited equivalents in each language and one would therefore not expect to find any substantial differences.

While there is evidence for the existence of common moral values across various cultures (Gibbs, Basinger, Grime, & Snarey, 2007), there might be factors which are common only to the cultures from which we have sampled our participants, thus limiting generalizability. For one, all three countries belong to developed westernized cultures. This might mean that participants from less developed nations, with non-western style cultures could show different reactions to judgments of other- and self-sacrifice (although it is worth noting that Sachdeva et al., 2015, failed to find any impact of culture—comparing US with Indian and Irani participants—on appropriateness of self-sacrifice in trolley dilemmas).

A possible limitation of our experiments is that we used only a few scenarios in the first study and only one in the second. Although we use a random effect for stimuli in analysis of Study 1, we were unable to reliably examine potential effects of scenarios on the judgment. When only a fixed set of stimuli is used in a study, the generalizability of its findings is always unknown (Bahník & Vranka, 2017) and stimuli sampling, when possible, is therefore recommended.

Sacrificial dilemmas used in psychological research come in many flavors (Christensen & Gomila, 2012). The dilemmas used in the present study largely shared features with the classical trolley dilemma. However, they necessarily differed in other features. For example, in all dilemmas, the actor was in a role of a worker who decided on the fate of his or her coworkers, which might have elicited a sense of responsibility. The results should therefore be interpreted in light of the limited breadth of situations depicted in the dilemmas as well as general limitations associated with the use of somewhat unrealistic sacrificial dilemmas (Bauman, McGraw, Bartels, & Warren, 2014; Kahane, 2015).

The goal of the present study was to explore how philosophical intuitions about moral judgments of self-sacrifice can be reconciled with conflicting empirical findings. For this purpose, we examined whether prescriptive or normative judgment leads to different self-other asymmetries. The results of three experiments where we manipulated the type of judgment, its framing, and the perspective from which the dilemmas were presented, showed that an other-serving bias emerges only for negatively framed normative judgments, while the results for other conditions were consistent with a self-serving bias or no bias at all. Future work could focus on identifying other instances where self-other asymmetries might be observed, how different judgment types affect these asymmetries, and whether these biases can be replicated in real-world settings.

Acknowledgements

The authors would like to thank Ira Theresa Maschmann for translating materials for Study 1a and administering it, Anita Körner for helpful discussion, Jana Dlouhá for creating a web application for Study 1b, and Dale Barr for statistical advice. The authors would like to especially thank Scio company for the data from the National Comparative Exams GAP test used in Study 2.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was partly supported by Deutsche Forschungsgemeinschaft (DFG-RTG 1253/2). The work of MV was supported by Progres Q18 grant by Charles University.

ORCID iD

Štěpán Bahník  <https://orcid.org/0000-0002-0579-6808>

Notes

1. Even though some deontologists might consider self-preservation a moral duty, the emphasis is on preserving one's moral, not biological life. Therefore, self-sacrifice is allowed in specific situations, unlike "self-murder" or harming others (Unna, 1998).
2. Methods, analyses, and hypotheses of Study 1 were pre-registered on Open Science Framework (<https://osf.io/fvd9t/>). Study 2 was pre-registered before any analysis was conducted but not before the data collection. Analysis scripts and data can be found at the same place. We did not conduct any other studies related to the three studies reported in the present paper. We did not run any a priori power analyses, but we report CIs for effect sizes, which show the uncertainty in estimates of the reported results. In Study 1, the sample size was limited by available funds and in Study 2, the sample size was limited by the number of applicants who took the admission test before which the questionnaire was administered.
3. The results of this analysis can be found on <https://osf.io/rnxp2/>.
4. Results for this analysis can be found on <https://osf.io/v9rgk/>.
5. In Study 1, it was not made explicitly clear whether "you" would survive if you sacrificed the other person. This possible confound was amended in Study 2.
6. The manner in which we analyze the data in Study 2 is slightly different from our preregistration. We separated the four questions by two factors (framing and judgment type) instead of treating them as a single factor with four levels as pre-registered. We considered this analysis to be more in line with the way the analysis was conducted in Study 1, thus facilitating comparisons between the studies. In the supplementary materials (<https://osf.io/9tj48/>), we do however report a table of results in the manner which we pre-registered.
7. Using data from Study 2, an analysis corresponding to a direct replication of Study 1 was conducted. It showed an interaction of judgment type and target of sacrifice in the same direction as in Study 1, which was however not significant (see <https://osf.io/2kpnh/> for full results).

References

- Andow, J. (2018). Are intuitions about moral relevance susceptible to framing effects? *Review of Philosophy and Psychology*, 9, 115–141.
- Bahník, Š., & Vranka, M. A. (2017). If it's difficult to pronounce, it might not be risky: The effect of fluency on judgment of risk does not generalize to new stimuli. *Psychological Science*, 28, 427–436.
- Barbosa, S., & Jiménez-Leal, W. (2017). It's not right but it's permitted: Wording effects in moral judgement. *Judgment and Decision Making*, 12, 308–313.
- Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, 8, 536–554.
- Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, 4, 353–376.
- Chisholm, R. M. (1963). Supererogation and offence: A conceptual scheme for ethics. *Ratio*, 5, 1–14.

- Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience & Biobehavioral Reviews*, *36*, 1249–1264.
- Cova, F., Strickland, B., Abatista, A., Allard, A., Andow, J., Attie, M., . . . , Cushman, F. (2018). Estimating the reproducibility of experimental philosophy. *Review of Philosophy and Psychology*. Advance online publication. doi:10.1007/s13164-018-0400-9
- Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*, 17320–17325.
- Di Nucci, E. (2013). Self-sacrifice and the trolley problem. *Philosophical Psychology*, *26*, 662–672.
- Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. *Psychology of Learning and Motivation*, *50*, 307–338.
- Elqayam, S., Thompson, V. A., Wilkinson, M. R., Evans, J. S. B., & Over, D. E. (2015). Deontic introduction: A theory of inference from is to ought. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1516.
- Elqayam, S., Wilkinson, M. R., Thompson, V. A., Over, D. E., & Evans, J. S. B. (2017). Utilitarian moral judgment exclusively coheres with inference from is to ought. *Frontiers in Psychology*, *8*, 1042.
- Epley, N., & Dunning, D. (2000). Feeling “holier than thou”: Are self-serving assessments produced by errors in self-or social prediction? *Journal of Personality and Social Psychology*, *79*, 861–875.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, *5*, 5–15.
- Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.
- Gibbs, J. C., Basinger, K. S., Grime, R. L., & Snarey, J. R. (2007). Moral judgment development across cultures: Revisiting Kohlberg’s universality claims. *Developmental Review*, *27*, 443–500.
- Goldstein, S. E., Tisak, M. S., & Boxer, P. (2002). Preschoolers’ normative and prescriptive judgments about relational and overt aggression. *Early Education and Development*, *13*, 23–40.
- Gray, K., & Wegner, D. M. (2011). Morality takes two: Dyadic morality and mind perception. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 109–127). Washington, DC: American Psychological Association.
- 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*, 364–371.
- Huebner, B., & Hauser, M. D. (2011). Moral judgments about altruistic self-sacrifice: When philosophical and folk intuitions clash. *Philosophical Psychology*, *24*, 73–94.
- Jacobs, R. A. (1987). Obligation, supererogation and self-sacrifice. *Philosophy*, *62*, 96–101.

- Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: Two faces of moral regulation. *Journal of Personality and Social Psychology, 96*, 521–539.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*, 54–69.
- Kahane, G. (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. *Social Neuroscience, 10*, 551–560.
- Lester, D. (2014). Altruistic suicide in the trolley problem. *Comprehensive Psychology, 3*, 6.
- Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin, 132*, 895–919.
- Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion, 14*, 573–587.
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science, 19*, 549–557.
- Nadelhoffer, T., & Feltz, A. (2008). The actor-observer bias and moral intuitions: Adding fuel to Sinnott-Armstrong's fire. *Neuroethics, 1*, 133–144.
- O'Hara, R. E., Sinnott-Armstrong, W., & Sinnott-Armstrong, N. A. (2010). Wording effects in moral judgments. *Judgment and Decision Making, 5*, 547–554.
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding mechanical Turk as a participant pool. *Current Directions in Psychological Science, 23*, 184–188.
- Pastötter, B., Gleixner, S., Neuhauser, T., & Bäuml, K. H. T. (2013). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition, 126*, 373–377.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28*, 369–381.
- Rohsenow, D. J. (1983). Drinking habits and expectancies about alcohol's effects for self versus others. *Journal of Consulting and Clinical Psychology, 51*, 752–756.
- Sachdeva, S., Iliev, R., Ekhtiari, H., & Dehghani, M. (2015). The role of self-sacrifice in moral dilemmas. *PLoS One, 10*(6), e0127409.
- Slote, M. (1984). Morality and self-other asymmetry. *The Journal of Philosophy, 81*, 179–192.
- Splawn, C. (2001). The self-other asymmetry and act-utilitarianism. *Utilitas, 13*, 323–333.
- Stocker, M. (1976). Agent and other: Against ethical universalism. *Australasian Journal of Philosophy, 54*, 206–220.
- Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology, 4*, 250.
- Unna, Y. (1998). Kant's argument against self-murder and its relation to the principle of self-preservation of reason (Unpublished doctoral dissertation). Boston University, Boston, MA.
- Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science, 18*, 689–690.

- Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44*, 1334–1338.
- Volz, L. J., Welborn, B. L., Gobel, M. S., Gazzaniga, M. S., & Grafton, S. T. (2017). Harm to self outweighs benefit to others in moral decision making. *Proceedings of the National Academy of Sciences, 114*, 7963–7968.
- Weinstein, N. D. (1987). Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample. *Journal of Behavioral Medicine, 10*, 481–500.
- Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*, 2020–2045.
- Wiegmann, A., Okan, Y., & Nagel, J. (2012). Order effects in moral judgment. *Philosophical Psychology, 25*, 813–836.

Author Biographies

Štěpán Bahník is a researcher at The Prague College of Psychosocial Studies. He is interested in judgment and decision making, moral judgment and behavior, and psychological methodology.

Emir Efendic is a postdoc at the University of Louvain's Psychological Sciences Research Institute in Belgium. He received his PhD from the University of Bordeaux and his work is focused on judgment and decision making.

Marek A. Vranka is a researcher at the Faculty of Social Sciences at Charles University in Prague. His main research topics are moral psychology, decision making and communication.