QJEP

Moral fatigue: The effects of cognitive fatigue on moral reasoning

Shane Timmons and Ruth MJ Byrne

Abstract

We report two experiments that show a moral fatigue effect: participants who are fatigued after they have carried out a tiring cognitive task make different moral judgements compared to participants who are not fatigued. Fatigued participants tend to judge that a moral violation is less permissible even though it would have a beneficial effect, such as killing one person to save the lives of five others. The moral fatigue effect occurs when people make a judgement that focuses on the harmful action, killing one person, but not when they make a judgement that focuses on the beneficial outcome, saving the lives of others, as shown in Experiment 1 ($n = 196$). It also occurs for judgements about morally good actions, such as jumping onto railway tracks to save a person who has fallen there, as shown in Experiment 2 ($n = 187$). The results have implications for alternative explanations of moral reasoning.

People make different decisions when they are tired, for example, judges make stricter parole decisions at the end of a decision session compared to at the start of a session (e.g., Danzigera, Levav, & Avnaim-Pesso, 2011; see also Mani, Mullainathan, Shafir, & Zhao, 2013; Spears, 2010). People may make different decisions when they are cognitively fatigued because their limited cognitive resources have been exhausted. They may no longer have sufficient capacity to allocate to new decisions (e.g., Baumeister & Heatherton, 1996; Schmeichel, Vohs, & Baumeister, 2003), or they may experience a reluctance to engage in further effortful processing (e.g., Inzlicht & Schmeichel, 2012). Their moral behaviour is also affected by the exhaustion of cognitive resources, for example, people are more inclined to cheat and deceive when they have carried out a task that is cognitively depleting, such as writing an essay with words that do not contain the letters "a" or "n" in it (e.g., Capraro & Cococcioni, 2016; Gino, Schweitzer, Mead, & Ariely, 2011; Mead, Baumeister, Gino, Schweitzer, & Ariely, 2009). We examine whether people experience decision fatigue for moral decisions about actions that violate a moral principle, such as harming one person, to bring about a beneficial outcome, such as saving many other people. Our novel aim is to test whether moral fatigue occurs because the exhaustion of cognitive resources affects people's ability to construct a model that

links actions to outcomes, and instead fatigued participants construct a simpler model that highlights the immoral nature of the action rather than its morally beneficial consequences.

We gave participants moral dilemmas of the following sort (from Moore, Clark, & Kane, 2008):

> You are the explosives expert for a company that has been hired to demolish a skyscraper. You are examining the last of the explosive charges when you notice a teenager below who is about to accidentally detonate one of the charges out of sequence. This explosion will result in the building's uncontrolled collapse onto you, the teenager, and the crowd of spectators. The teenager is several floors below you and cannot hear you because of the loud demolition noise. You realize that the only way to stop the teenager from detonating the charge is to flip a switch that reactivates the building's electricity. Because he is touching an open circuit, this will electrocute him but will prevent the explosion.

School of Psychology and Institute of Neuroscience, Trinity College Dublin, University of Dublin, Dublin, Ireland

Corresponding author:
Shane Timmons, School of Psychology and Institute of Neuroscience, Trinity College Dublin, University of Dublin, Dublin 2, Ireland.
Email: timmonss@tcd.ie

When people are asked whether it is permissible to kill the teenager, some people's judgements appear to reflect the deontological view that an action such as killing someone is a morally wrong violation of a core principle, whereas other people's judgements appear to reflect the utilitarian principle that a violation that leads to an outcome that benefits many people is justified (e.g., Baron, 2017; Białek & De Neys, 2017; Bonnefon, Shariff, & Rahwan, 2016; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). We examine whether cognitive fatigue affects people's deontological and utilitarian judgements. People must allocate cognitive resources to weigh up the benefits of the outcome against the moral violation in the action when they make the utilitarian judgement that the action is permitted. Their judgements about such dilemmas require them to consider both the action and the outcome and how they are linked (e.g., Wiegmann & Waldmann, 2014). The utilitarian decision may depend on constructing a model that makes explicit the causal links between the otherwise immoral action, killing a person and its outcome, saving several other people (e.g., Crockett, 2013; Cushman, 2013; see also Lagnado, Gerstenberg, & Zultan, 2013; Parkinson & Byrne, 2017a, 2017b). The link between the action and the outcome provides a justification or reason for the action. People require sufficient cognitive resources to be able to simulate both components and their relations. Conversely, when people make the deontological decision that the action is not permitted, they may have evaluated the action in isolation from its outcome (e.g., Patil, 2015). The condemnation of the action may arise from representing the experience of performing the action, rather than the experience of the outcome (e.g., Miller, Hannikainen, & Cushman, 2014). People judge the action should not be taken when they mentally represent it vividly (e.g., Amit & Greene, 2012). Similarly, their moral judgements are affected when they focus on the actor rather than the recipient (e.g., Grey, Waytz, & Young, 2012; Grey & Wegner, 2009). And people remain averse to harmful actions even when the causal link to an outcome is removed, such as shooting a person with a fake gun (e.g., Cushman, Grey, Gaffey, & Mendes, 2012). Although people seem insensitive to the outcome when a moral judgement highlights the required action, such as harm caused to a protected value, their judgements can be changed by a focus on the beneficial outcome, such as the net benefits for the value (e.g., Bartels & Medin, 2007). Hence, we consider that cognitive fatigue may affect people's ability to construct a more complex model that links the action and the outcome, for example, to link the action of killing a person to the outcome of saving others. We test the idea that when participants are cognitively fatigued, they will be less able to construct such a model to reason about a dilemma, and so if they have constructed a model that focuses only on the immoral action, killing a person, they will be more inclined to judge that the action is not permitted.

Cognitive fatigue

We expect that a moral fatigue effect should occur if moral judgements depend at least in part on cognitive reasoning processes (e.g., Bucciarelli & Daniele, 2015; Bucciarelli, Khemlani & Johnson-Laird, 2008; Kohlberg, 1976; Piaget, 1932), rather than solely on automatic, emotional, or intuitive reactions (e.g., Damasio, 2000; Haidt, 2001; see also Pizarro & Salovey, 2002; Rozin, Lowery, Imada, & Haidt, 1999). People have limited abilities to carry out executive functions, such as allocating attention, manipulating information in working memory, and inhibiting prepotent responses (e.g., Baddeley, 1996, 2007; Smith & Jonides, 1999). One way to examine whether people rely on effortful reasoning processes to make moral judgements is to exploit their limited capacity, for example, to test their judgements under working memory load. The logic of dual-task designs is to rely on a simultaneous secondary task, to tax cognitive resources in parallel. Secondary task loads compromise executive functioning by dividing attention resources between the primary and secondary tasks (e.g., Gilbert & Hixon, 1991; Lavie, Hirst, De Fockert, & Viding, 2004; Ward & Mann, 2000). And secondary tasks have been found to affect moral reasoning, for example, the decision to violate a moral principle to bring about a greater good takes longer to make when the decision is made under conditions of cognitive load (e.g., Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). Moreover brain regions associated with cognitive control have been implicated in moral judgements (e.g., Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene et al., 2001). The effects of secondary tasks on moral judgement have been taken to indicate that utilitarian moral judgements depend on controlled reasoning processes (e.g., Conway & Gawronski, 2013; Greene et al., 2008; Trémolière, De Neys, & Bonnefon, 2012), although it has also been argued that utilitarian and deontological judgements could both be rooted in intuition (e.g., Białek & De Neys, 2017; Landy & Royzman, 2018). Note that secondary tasks compete for cognitive resources and so their effects are different from cognitive tasks that encourage reasoning, for example, when participants carry out cognitive tasks that require deliberative thought such as those that comprise the cognitive reflection test, they subsequently make *more* utilitarian judgements, presumably because the prior cognitive tasks encourage controlled reasoning (e.g., Paxton, Ungar, & Greene, 2012; Yilmaz & Saribay, 2017). In contrast, secondary tasks compete for cognitive resources and lead to fewer utilitarian judgements. Analogously, the logic of a *sequential* task design is to exhaust cognitive resources by employing a sequential temporal load, that is, participants first carry out a cognitively exhausting task, and then immediately afterwards they engage in some higher-order cognitive task. Executive functions draw upon the same resource and when this resource becomes exhausted, people's ability to engage in higher-order cognitive processes

becomes impaired (e.g., Baumeister & Heatherton, 1996; Muraven & Baumeister, 2000; Schmeichel et al., 2003). For example, tasks that involve reasoning, cognitive extrapolation, and thoughtful reading comprehension are impaired when participants are cognitively fatigued, whereas less complex tasks, such as general knowledge tests and simple recall tests, are unaffected (e.g., Schmeichel, 2007). Hence, we aim to test whether people rely on reasoning to make moral judgements, by examining the effect of sequential cognitive depletion tasks on their moral judgements.

Reservations have been expressed about the phenomenon of depletion, in particular, about the effect size of sequential task-induced cognitive fatigue, which can be very small, at least for depleting tasks which participants do not experience as cognitively effortful or for depleting tasks that are demanding but not based on breaking a habit (see Baumeister & Vohs, 2016; Carter & McCullough, 2014; Dang, 2016; Hagger et al., 2016). In contrast, sequential task-induced cognitive fatigue appears to be robust in depleting tasks for which participants have formed a habit, such as writing essays, when they must do so without using the letters "a" and "n," or re-typing a paragraph, when they must do so without using the letter "e" or the spacebar (e.g., Hagger, Wood, Stiff, & Chatzisarantis, 2010; Muraven, Pogarsky, & Shmueli, 2006; Schmeichel, 2007). Our aim in the experiments we report is not to test claims made about the nature of depletion but rather to use the sequential load method of depletion studies, analogous to the simultaneous load method of working memory studies, to reduce reliance on cognitive resources. Our aim is to test whether people make different moral judgements when they are cognitively fatigued, specifically, whether people who are fatigued tend to judge that an action such as killing a person to save others, is less permissible compared to people who are not fatigued. We aim to examine whether differences in their judgements arise because they have constructed different sorts of models of the relation between an action and its outcome.

Experiment 1

The aim of the experiment was to examine whether a cognitive fatigue effect occurs for moral judgements because people construct a simple model that fails to explicitly link the action to the outcome when they are fatigued. Hence, we test whether participants who are fatigued make different judgements compared to participants who are not fatigued for judgements that focus on the action and judgements that focus on the outcome. We expect to observe a moral fatigue effect when the judgement explicitly mentions the action:

Killing the teenager in this case is morally . . .

That is, we expect that participants who are fatigued will tend to judge that the action is less permissible compared to those who are not fatigued. However, we expect that when participants' attention is explicitly directed to the outcome, even those who are fatigued will construct a model that links the action to the outcome and tend to judge that the action is more permissible compared to when their attention is not directed to the outcome:

Doing this in order to save yourself and the crowd of spectators is morally . . .

Hence we expect the moral fatigue effect will be diminished when participants make judgements that explicitly mention the outcome, compared to judgements that explicitly mention the action.

We manipulated one other factor primarily as a control. Many studies distinguish moral judgements about "impersonal" dilemmas in which the physical action is indirect, such as killing someone by flipping a switch to reactivate the building's electricity, and emotive "personal" dilemmas in which the physical action is more direct:

You realize that the only way to stop the teenager from detonating the charge is to drop a heavy cinderblock on his head. This will crush his skull and kill him almost instantly but will prevent the out-of-sequence explosion.

Participants tend to judge that the action is not permitted in "personal" dilemmas and they tend to judge it is permitted in "impersonal" ones (e.g., Greene et al., 2001; Mikhail, 2007; Nichols & Mallon, 2006). We included personal and impersonal dilemmas merely to check whether any differences between fatigued and non-fatigued participants for action-focused and outcome-focused judgements occurred for the two sorts of dilemma. The personal and impersonal versions of the dilemmas differed only in the directness of killing, and controlled for potential confounds such as phrasing, number of deaths, and word length (from Moore et al., 2008; see also Paxton et al., 2012).

The participants' task was to make the following sort of judgement:

Killing the teenager in this case is morally:

1____2____3____4____5____6____7
Forbidden   Obligatory

We chose the first-person perspective and asked for a normative judgement, rather than a predictive response such as "would you do it?" to control for potential confounds (e.g., Amit & Greene, 2012; Valdesolo & DeSteno, 2006; cf. Cushman, Knobe, Sinnott-Armstrong, 2008). There is considerable variation in the measures used in studies of moral judgement, which can make comparisons across studies difficult. Measures differ in their formats, from

forced-choice, dichotomous measures (e.g., Amit & Greene, 2012; Bucciarelli et al., 2008; Conway & Gawronski, 2013; Cushman, Sheketoff, Wharton, & Carey, 2013), to Likert-type scales (e.g., Bartels, 2008; Cushman, Young, & Hauser, 2006; Lombrozo, 2009), or both (e.g., Cushman et al., 2012; Paxton et al., 2012). There is also diversity in the type of normative judgement asked about the action, such as whether it is appropriate (Greene et al., 2008; Moore et al., 2008), wrong (Graham, Haidt, & Nosek, 2009; Laham, Alter, & Goodwin, 2009), acceptable (Bartels, 2008; Greene et al., 2009), permissible (Lombrozo, 2009; Ugazio, Lamm, & Singer, 2012), ethical (Paharia, Kassam, Greene, & Bazerman, 2009), or obligatory (Cushman et al., 2006; O' Hara, Sinnott-Armstrong & Sinnott-Armstrong, 2010). We chose a scaled response format, from forbidden to obligatory, with permissible as the implicit mid-point, rather than a dichotomous format, to allow a more nuanced response in that participants could indicate that an action was not permissible, or that it was permissible (but not necessarily obligatory), or that it was obligatory (e.g., Kahane & Shackel, 2010; see also Cushman et al., 2006; Verschueren, Schaeken, & d'Ydewalle, 2005). A scale that permits judgements not only of permissibility and impermissibility but also of obligation also allows the comparison of judgements about morally bad actions, examined in Experiment 1, and judgements about morally good actions, examined in Experiment 2. Moreover, for judgements about morally bad actions, some people consider actions such as harming one person to save others to be a permissible choice of a decision maker, rather than an obligatory duty, exhibiting a "moral minimalism," but others judge such actions to be obligatory, exhibiting a "strict utilitarianism" (e.g., Royzman, Landy, & Leeman, 2015). Similarly, for judgements about morally good actions, some people may think of actions such as carrying out a self-sacrificial action to save another person, as an obligatory duty whereas others may consider it merely a permissible choice (e.g., Algoe & Haidt, 2009). Our scale of forbidden through permissible to obligatory enables a more complete assessment of participants' judgements.

We examined not only participants' moral judgements but also how they felt about their moral decisions. The role of emotion in moral judgement remains controversial (e.g., Kahane, Everett, Earp, Farias, & Savulescu, 2015; Koenigs, Kruepke, Zeier, & Newman, 2012; Valdesolo & DeSteno, 2006; but see Landy & Goodwin, 2015). We examine emotion as a consequence of moral judgement, since depletion can affect emotion regulation (e.g., Baumeister, Vohs, & Tice, 2007; Hofmann, Rauch, & Gawronski, 2007; Johns, Inzlicht, & Schmader, 2008). Of course, people may anticipate how they will feel as a consequence of a moral decision and their anticipation may in turn affect the decision they make (e.g., Tasso, Sarlo, & Lotto, 2017).

Method

Participants. The participants were 196 individuals who completed the experiment on two online platforms, CrowdFlower and Prolific Academic. A further 28 participants were eliminated prior to analysis because English was not their first language ($n = 4$), they had duplicate IP addresses ($n = 2$), or they failed to carry out the instructions in the writing task to re-write the presented paragraph and not to type the letter "e" or use the spacebar key ($n = 22$). There were 128 women and 63 men, 4 people who indicated their gender as other than male or female and 1 who indicated a preference not to say, and the average age was 33 years with a range from 18 to 69 years. We restricted participation to a set of countries that had English as a first language and so most of the participants were from the United States ($n = 103$), the United Kingdom ($n = 78$), Ireland ($n = 8$), Australia ($n = 4$), New Zealand ($n = 2$), and Canada ($n = 1$). Participants received a nominal payment in line with their platform norms; 25 cents (US$) on Crowd-Flower and £1.50 (GBP) on Prolific Academic. They were assigned at random to one of four groups: fatigued-outcome ($n = 41$), fatigued-action ($n = 51$), non-fatigued-outcome ($n = 58$), and non-fatigued-action ($n = 46$). Sample size was initially calculated on the basis of a moderate to large effect size in laboratory-based cognitive depletion in most published studies and a high correlation between the repeated measures of personal and impersonal dilemmas (e.g., Hagger et al., 2010), that is, approximately 20 participants per cell. However, following comments on an earlier draft, sample size was subsequently reset to approximately 50 participants per cell in line with recommendations in Simmons, Nelson, and Simonsohn (2011) and further participants were recruited; in fact, the recalculated sample size made no difference to the results.

Materials and design. The design was a 2 (fatigue: fatigued vs. non-fatigued) × 2 (dilemma: personal vs. impersonal) × 2 (judgement: action vs. outcome) design, with repeated measures on the second factor. Participants were given four moral dilemmas, two personal and two impersonal, in randomised order (see the Supplementary material). We used four different contents for the moral dilemmas and assigned the contents at random to the personal and impersonal versions in two ways to create two sets, to control for content effects, and each participant received one set at random. For each dilemma, they were asked to make a moral judgement, for example,

Killing the teenager in this case is morally:

1_____2_____3_____4_____5_____6_____7
Forbidden   Obligatory

They made their moral judgement on a scale from 1 (*forbidden*) to 7 (*obligatory*). Half of the participants were
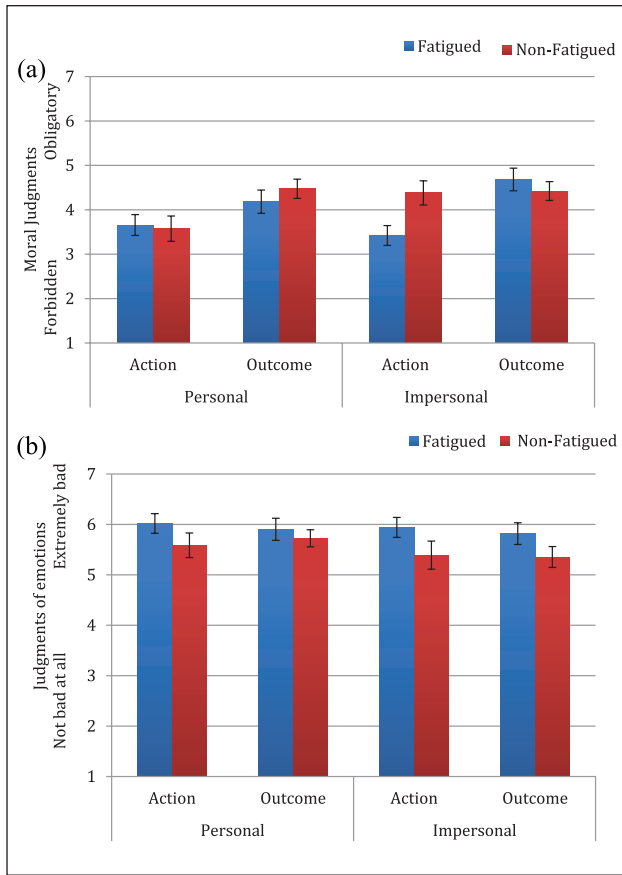
Figure 1. (a) Mean moral judgements and (b) mean emotion judgements, for morally bad dilemmas in Experiment 1. Error bars are standard error of the mean.

given the judgement framed to highlight the action, for example, "Killing the teenager in this case is morally . . ." and the other half were given the judgement framed to highlight the outcome, for example, "Doing this in order to save yourself and the crowd of spectators is morally . . ." Participants were also asked "how bad would this decision make you feel?" They rated how they felt about their decision from 1 (*not bad at all*) to 7 (*extremely bad*).

Participants completed an online depletion task (adapted from Muraven et al., 2006). They were asked to re-type one 150-word paragraph taken from a statistics book as quickly as possible. Then they were asked to re-type a second paragraph (see the Supplementary material). Participants in the fatigued group were told they were not to type the letter "e" or use the spacebar key, thus breaking a previously formed typing habit. Participants in the non-fatigued group were given no constraints. Participants rated the difficulty of the re-typing task, on a scale from 1 (*not at all difficult*) to 7 (*extremely difficult*), to determine whether it was sufficiently effortful, which is an important manipulation check for sequential task designs (see Dang, 2016). Other manipulation checks included the Brief Mood Introspection Scale (e.g., Mayer & Gaschke, 1988; see Schmeichel, 2007;

Valdesolo & DeSteno, 2006), and they also rated the difficulty of each of the other tasks on a scale from 1 (*not at all difficult*) to 7 (*extremely difficult*), and the results are provided in the Supplementary material.

Procedure. The materials were presented using Survey-Gizmo software, presented on CrowdFlower or Prolific Academic to recruit participants. Each dilemma was presented on a single screen with the scale below it. The other tasks were presented on separate screens. The experiment took approximately 20 min to complete.

Results and discussion

The raw data files for both experiments are available at: https://reasoningandimagination.wordpress.com/data-archive/

The manipulation checks confirmed that participants in the fatigue conditions rated their typing task (*Mdn* = 5, interquartile range [*IQR*] = 5-6) as more difficult than participants in the non-fatigue conditions (*Mdn* = 3, *IQR* = 2-5), Mann–Whitney $U = 2,590.5$, $p < .001$, $r = .40$ (we provide medians and interquartile ranges for the manipulation checks because the data are ordinal based on single response Likert-type scales). They also rated the moral judgement task as more difficult (*Mdn* = 4, *IQR* = 3-6) compared to the non-fatigued participants (*Mdn* = 4, *IQR* = 1.25-5), $U = 4,004$, $p = .046$, $r = .14$ (for further details, see the Supplementary material).

Participants tended to judge the actions to be permissible, with mean judgements of 4 on the 1 to 7 scale (in which 1 is *forbidden*, 7 is *obligatory*, and the mid-point 4 implicitly is *permissible*), as Figure 1a shows. Responses to personal and impersonal dilemmas were approximately normally distributed around the mean of 4 (skewness = −0.23 and −0.27; kurtosis = −0.88 and −0.86, respectively). We carried out a 2 (fatigue: fatigued, non-fatigued) × 2 (dilemma: personal, impersonal) × 2 (judgement focus: outcome, action) analysis of variance (ANOVA) with repeated measures on the second factor, on the moral judgements. The results showed that the three factors interacted, $F(1, 192) = 13.64$, $p < .001$, $\eta_p^2 = .07$, a medium effect size as Figure 1a shows. There was no main effect of fatigue, $F(1, 192) = 1.08$, $p = .299$, $\eta_p^2 = .01$; there was a main effect of judgement focus, $F(1, 192) = 9.61$, $p = .002$, $\eta_p^2 = .05$, as participants tended to judge that the action was less permissible when the judgement focused on the action rather than the outcome; and a main effect of dilemma, $F(1, 192) = 5.75$, $p = .017$, $\eta_p^2 = .03$, as participants tended to judge that the action was less permissible in personal dilemmas than impersonal ones. Fatigue did not interact with judgement focus, $F(1, 192) = 0.94$, $p = .333$, $\eta_p^2 = .01$, or dilemma type, $F(1, 192) = 1.25$, $p = .266$, $\eta_p^2 = .01$, and judgement focus did not interact with dilemma type, $F(1, 192) = 0.93$, $p = .761$, $\eta_p^2 = .001$.

We decomposed the significant three-way interaction to test our hypotheses about the expected differences between participants in the fatigued and the non-fatigued conditions and between action-focused and outcome-focused judgements, with a Bonferroni corrected alpha of .006 for the eight key comparisons. We expected to observe effects of fatigue for action-focused judgements, and the three-way interaction arises largely because such effects were indeed observed, but for impersonal dilemmas and not for personal ones, as Figure 1a shows. Fatigued participants judged actions in action-focused impersonal dilemmas to be less permissible compared to non-fatigued participants, somewhat marginally so on the corrected alpha, $t(95) = 2.74$, $p = .007$, $d = 0.56$; there were no other differences between fatigued and non-fatigued participants: action-focused personal, $t(95) = 0.19$, $p = .847$, $d = 0.04$; outcome-focused impersonal, $t(97) = 0.79$, $p = .434$, $d = 0.16$; and outcome-focused personal, $t(97) = 0.86$, $p = .392$, $d = 0.17$. Actions in action-focused judgements were judged less permissible than actions in outcome-focused judgements by fatigued participants in impersonal dilemmas, $t(90) = 3.73$, $p < .001$, $d = 0.79$; there were no other significant differences on the corrected alpha of .006 between action- and outcome-focused judgements: fatigued personal, $t(90) = 1.52$, $p = .131$, $d = 0.32$; non-fatigued personal, $t(102) = 2.57$, $p = .012$, $d = 0.51$; and non-fatigued impersonal, $t(102) = 0.12$, $p = .902$, $d = 0.02$.

Although our hypotheses did not concern the personal and impersonal factor, we note for completeness that non-fatigued participants tended to show the well-documented effect of judging that the action was less permissible for personal dilemmas than impersonal ones; the difference occurred for action-focused judgements, $t(45) = -3.99$, $p < .001$, $d = 0.59$, but not for outcome-focused ones, $t(57) = 0.25$, $p = .802$, $d = 0.03$; fatigued participants showed no effects for action-focused, $t(50) = 1.13$, $p = .263$, $d = 0.16$, or outcome-focused judgements, $t(40) = 2.04$, $p = .049$, $d = 0.32$.

Participants indicated that they felt bad about their moral judgements, an average of about 5.5 on the 1 to 7 scale in which 7 = extremely bad. An ANOVA of the same design on how participants felt about their judgements showed a main effect of fatigue, $F(1, 192) = 4.12$, $p = .044$, $\eta_p^2 = .02$, as fatigued participants reported feeling worse about their judgements; no main effect of judgement focus, $F(1, 192) = 0.03$, $p = .859$, $\eta_p^2 < .001$; and a main effect of dilemma type, $F(1, 192) = 4.74$, $p = .031$, $\eta_p^2 = .02$, as participants felt worse about personal compared to impersonal dilemmas; and no interactions of the variables: fatigue and dilemma, $F(1, 192) = 1.44$, $p = .231$; fatigue and judgement focus, $F(1, 192) = 0.18$, $p = .671$, $\eta_p^2 = .001$; judgement focus and dilemma, $F(1, 192) = 0.30$, $p = .588$, $\eta_p^2 = .002$; and fatigue and judgement focus and dilemma, $F(1, 192) = 0.25$, $p = .617$, $\eta_p^2 = .001$, as Figure 1b shows.

The results show a moral fatigue effect for judgements about morally bad actions: participants who were fatigued judged that a bad action, such as killing a teenager by flipping a switch to reactivate a building's electricity, was less permissible compared to participants who were not fatigued, when the judgement directed their attention to the action but not when it directed their attention to the outcome; an effect that occurs only for impersonal dilemmas. For personal dilemmas, the frequently observed and robust tendency for participants to judge that the morally bad action, such as dropping a cinderblock on the teenager's head, is impermissible tends to overshadow any effects of fatigue.

The expected two-way interaction of fatigue and judgement focus occurs for impersonal dilemmas but not for personal ones, and hence fatigued participants tend to judge the action to be as impermissible for impersonal dilemmas as for personal ones, and so they do not discriminate between personal and impersonal dilemmas in the way that non-fatigued participants do. When their attention is directed to the outcome in the outcome-focused judgements, they make similar judgements to non-fatigued participants. The result corroborates the idea that participants who have engaged in a cognitively tiring task tend to judge that the harmful action is less permissible than participants who have engaged in a less tiring task because they construct a simpler model of the events that does not explicitly link the harmful action to its beneficial outcome. When their attention is explicitly directed to the outcome, however, they overcome this limitation.

Participants tended to judge the actions to be permissible (an average of 4 on the 1-7 scale), and we have described ratings of less than 4 as "less permissible" here. It could be argued that a rating of "3" or "2" is intended instead to indicate "forbidden" rather than "less permissible." However, it seems plausible that a participant who wished to indicate a judgement of "forbidden" would choose "1," which was labelled "forbidden."

The dilemmas used in the experiment have been widely used (e.g., Greene et al., 2001; Gürçay & Baron, 2017; Moore et al., 2008). Their content differs in important ways, such as the number of individuals to be saved, the relationship of the actor to the individual to be harmed, and whether the actor's own life is to be saved, and so the moral fatigue effect is not restricted to a particular sort of dilemma (see Supplementary Material). However, we note that the vaccine dilemma, although widely used, may be somewhat flawed: participants may believe they could determine which substance is the vaccine and which is the lethal one by testing only one substance, rather than both, and so there would be only a 50% risk of killing a person. Nonetheless, properties of a single dilemma cannot account for the differences we observed in the experiment, since participants in every condition received the same dilemmas. The next experiment examines whether the

moral fatigue effect occurs when people reason about morally good actions, such as the noble self-sacrificial deeds that lead to the experience of moral elevation.

Experiment 2

The aim of the experiment was to examine whether the cognitive fatigue effects observed for judgements about moral violations extend to judgements about morally good deeds, for judgements that focus on actions, and not for judgements that focus on outcomes. People are uplifted and inspired when they witness or read about acts of moral goodness, noble or self-sacrificial actions, such as a man jumping on the railway tracks to lie on top of another man who has fallen there, to save him from an oncoming train (e.g., Algoe & Haidt, 2009; Freeman, Aquino, & McFerran, 2009; Lai, Haidt, & Nosek, 2014). People often wish to emulate such moral goodness when they experience moral elevation (e.g., Algoe & Haidt, 2009; Cox, 2010; Schnall, Roper, & Fessler, 2010). Comparatively few studies have examined the cognitive processes underlying reasoning about morally good actions (for a review, see Pohling & Diessner, 2016). We test the idea that when people make judgements about whether such morally elevating acts should be taken, they must also construct a model in which they link the self-sacrificial act to the beneficial outcome. Hence, we predict that moral fatigue effects will occur even when people reason about self-sacrificial morally good actions.

We used the same design as the previous experiment to examine whether individuals who were fatigued made different moral judgements about these good actions. Our interest once again is in the interaction of fatigue with judgement focus, and we examine whether participants who are fatigued judge that an action such as jumping onto the railway tracks is less obligatory when the judgement focuses on the action rather than the outcome. For comparison with the previous experiment, we also include personal and impersonal self-sacrificial dilemmas. There has hitherto been no examination of whether people make different judgements about self-sacrificial dilemmas that are personal or impersonal and it is unknown whether it is a dimension of relevance for moral judgements about good actions. We created personal and impersonal versions of real newspaper stories, for example, in the personal version, the man jumped down on the tracks and laid on top of the person who had fallen there, whereas in the impersonal version, the man jumped down on the tracks and pulled a lever to divert the train onto another track away from the person who had fallen there. We framed the judgements to focus on the action, for example, "In your opinion, Mr Autrey jumping in front of the train in this case was morally . . ." or to focus on the outcome, for example, "In your opinion, doing this to save Mr Hollopeter was morally . . ."

Method

Participants. The participants were 187 volunteers who completed the study on the online platforms CrowdFlower and Prolific Academic. Prior to any data analysis, a further 6 participants were removed as English was not their first language and 19 were removed for failing to follow the instructions on the writing task. The participants were 115 women and 69 men and 3 participants reported their gender as other. Their average age was 35 years with a range from 18 to 72 years old. The participants were from the United States ($n=101$), the United Kingdom ($n=77$), Australia ($n=4$), Ireland ($n=2$), New Zealand ($n=1$), Canada ($n=1$), and one American participant in Venezuela. Participants received 25 cents (US$) on CrowdFlower and £1.50 (GBP) on Prolific Academic. They were assigned at random to one of four groups: fatigued-outcome ($n=46$), fatigued-action ($n=46$), non-fatigued-outcome ($n=46$), and non-fatigued-action ($n=49$). Sample size was calculated in the same way as the previous experiment.

Materials, design, and procedure. The design and procedure was the same as the previous experiment. The materials were two newspaper articles in their original form, as well as two modifications of them to create impersonal versions (see the Supplementary material). Participants read one personal and one impersonal story, and they received one version of each of the stories (i.e., either Subway-Personal and Baseball-Impersonal or Subway-Impersonal and Baseball-Personal). The stories were presented in a different randomised order for each participant. Participants made the same moral judgements as the previous experiment using the same scale from 1 (*forbidden*) to 7 (*obligatory*), they also judged how they felt about their decision in the same way as the previous experiment, and the depletion task was the same as the previous experiment.

Participants completed several manipulation checks including the mood scale and difficulty ratings used in the previous experiment. They also completed a shortened moral elevation scale to check that the stories were morally inspiring: they were asked to indicate how much they experienced or were still experiencing the following emotions or thoughts while reading the story (on a 1-7 scale where 1 = *not at all* and 7 = *a lot*): (1) inspired, (2) there is still some good in the world, and (3) the person in the story has shown me how to be a better person, and the results are provided in the Supplementary material. They completed the tasks in the following order: fatigue task, mood scale, moral elevation judgement, moral judgement, emotion judgement, and difficulty ratings.

Results and discussion

The manipulation checks confirmed that participants in the fatigue conditions rated their writing task as significantly
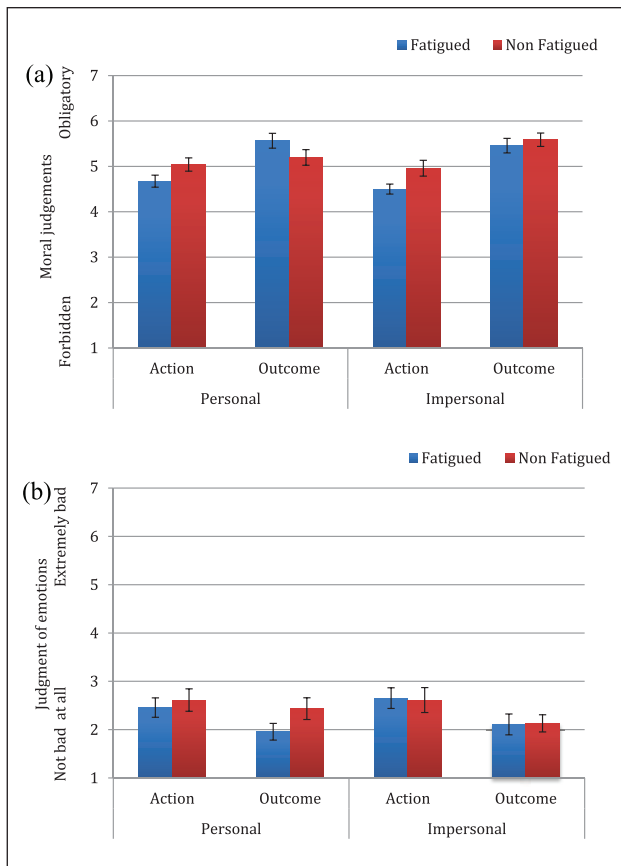
Figure 2. (a) Mean moral judgements and (b) mean emotion judgements, for morally good dilemmas in Experiment 2. Error bars are the standard error of the mean.

more difficult ($Mdn=5$, $IQR=5$-6) than those in the non-fatigue groups ($Mdn=4$, $IQR=2$-5), $U=2,538$, $p<.001$, $r=.54$; they did not differ in their ratings of the difficulty of the moral judgement task for the fatigue ($Mdn=2$, $IQR=1$-3) and non-fatigue conditions ($Mdn=2$, $IQR=1$-3), $U=4,034$, $p=.341$.

Participants tended to judge the actions to be somewhat obligatory, with mean judgements of 5 on the 1 to 7 scale (in which 7 is obligatory), as Figure 2a shows. Responses to personal and impersonal stories were approximately normally distributed around the mean of 5 (skewness$=-0.29$ and -0.03; kurtosis$=0.61$ and -0.39, respectively). An ANOVA of the same design as the previous experiment on moral judgements showed once again no main effect of fatigue, $F(1, 183)=1.33$, $p=.250$, $\eta_p^2 =.01$, and once again a main effect of judgement focus, $F(1, 183)=26.72$, $p<.001$, $\eta_p^2 =.13$, as participants tended to judge the action to be less obligatory for judgements that focused on the action rather than the outcome, and this time there was no main effect of dilemma, $F(1, 183)=0.01$, $p=.936$, $\eta_p^2 <.001$. The key two-way interaction of fatigue and judgement focus was significant, $F(1, 183)=4.38$, $p=.038$, $\eta_p^2 =.023$, a small effect size; fatigue and dilemma

did not interact, $F(1, 183)=3.08$, $p=.081$, $\eta_p^2 =.02$, nor did judgement focus and dilemma, $F(1, 183)=2.54$, $p=.113$, $\eta_p^2 =.01$, and the three variables did not interact, $F(1, 183)=1.46$, $p=.229$, as Figure 2a shows.

The decomposition of the two-way interaction of fatigue and judgement focus with a Bonferroni correction of .0125 for four comparisons shows that fatigued participants tended to judge the action to be less obligatory for action-focused judgements than outcome-focused ones, $t(90)=5.50$, $p<.001$, $d=1.16$; there was no difference for the non-fatigued participants on the corrected alpha of $p<.0125$, $t(93)=2.06$, $p=.043$, $d=0.43$. Fatigued participants judged the action to be marginally less obligatory than non-fatigued participants for action-focused judgements on the corrected alpha of $p<.0125$, $t(86.13)=2.40$, $p=.018$, $d=0.49$; there were no differences between the groups for outcome-focused judgements, $t(90)=0.64$, $p=.525$, $d=0.13$. This two-way interaction of fatigue and judgement focus for morally good actions is consistent with the interaction of fatigue and judgement focus for morally bad actions observed in the previous experiment, for impersonal dilemmas. We note that the personal and impersonal nature of the dilemmas showed no main effect and did not interact with any other variable in this experiment, and we tentatively suggest that this factor may not be as influential for judgements about morally good actions as it is for morally bad actions.

Participants indicated that they did not feel bad about their moral judgements, an average of about 2 on the 1 to 7 scale in which 1$=$not bad. An ANOVA of the same design as the previous one on the emotion ratings showed that unlike the previous experiment, there was no main effect of fatigue, $F(1, 183)=0.68$, $p=.409$, $\eta_p^2 =.004$, a main effect of judgement focus, $F(1, 183)=5.23$, $p=.023$, $\eta_p^2 =.03$, as participants did not feel as good when they made judgements that focused on the action compared to judgements that focused on the outcome, as Figure 2b shows, and no main effect of dilemma, $F(1, 183)=0.10$, $p=.919$, $\eta_p^2 <.001$. There was no interaction between fatigue and dilemma, $F(1, 183)=2.33$, $p=.129$, $\eta_p^2 =.01$, or judgement focus, $F(1, 183)=0.27$, $p=.607$, $\eta_p^2 =.001$; dilemma did not interact with judgement focus, $F(1, 183)=0.66$, $p=.417$, $\eta_p^2 =.004$; and there was no three-way interaction, $F(1, 183)=0.37$, $p=.543$, $\eta_p^2 =.002$. The lack of effects of fatigue on emotion judgements for self-sacrificial actions may suggest that these good actions require less justification for participants to evaluate them positively.

The experiment shows a moral fatigue effect for judgements about morally elevating actions—fatigued participants judged morally good actions, such as jumping on to the railway tracks, to be less obligatory when the judgement focused on the self-sacrificial action compared to when it focused on the beneficial outcome, saving a person who had fallen there; there was no effect for non-fatigued

participants. The result is consistent with the finding of the previous experiment in which fatigued participants judged morally bad actions, such as flipping a switch that would electrocute a teenager, to be less permissible when the judgement focused on the bad action compared to when it focused on the beneficial outcome, saving many others; there was no effect for non-fatigued participants. The difference between the two experiments is that the interaction of fatigue and judgement focus for morally bad actions occurred only for impersonal dilemmas, whereas for morally good actions, it occurred for both personal and impersonal dilemmas.

The results were observed using a scale that ranged from "forbidden" to "obligatory," with an implicit midpoint of "permissible," which we have suggested enables a more complete assessment of judgements suited for testing morally good outcomes as well as morally bad ones. The results of Experiments 1 and 2 suggest that it performed as expected. In any case, the nature of the scale does not modify the interpretation of the results, since the same scale was used in each condition in the experiments.

The results of the experiment again corroborate the idea that participants who have engaged in a cognitively tiring task construct a model of the events that does not explicitly link the action to its beneficial outcome, whether it is a morally good self-sacrificial action, or an action that violates a moral principle. When their attention is directed to the outcome, they overcome this limitation.

General discussion

Participants who have completed a cognitively tiring task tend to judge that a harmful action, such as killing a person, that leads to a good outcome, saving several others, is less permissible compared to participants who have completed a less cognitively tiring task. The moral fatigue effect occurs for judgements that focus on the harmful action but not for judgements that focus on the beneficial outcome: When their attention is directed to the outcome, fatigued and non-fatigued participants make similar judgements, as Experiment 1 shows. The result corroborates the idea that participants who have engaged in a cognitively tiring task judge that the harmful action is not permitted because they construct a simple model of the events that does not explicitly link the harmful action to its beneficial outcome. When their attention is directed to the outcome, they overcome this limitation. The effect occurs only for impersonal dilemmas—fatigued participants tend to judge that the action is less permissible for impersonal dilemmas just as much as for personal ones, and so they do not discriminate between personal and impersonal dilemmas in the way that non-fatigued participants do. Participants also show a moral fatigue effect for judgements about self-sacrificial good deeds. Participants who have completed a cognitively tiring task tend to judge

that a helpful action that leads to a good outcome, such as jumping on to the railway tracks to save a person who has fallen there, is less obligatory compared to participants who have completed a less cognitively tiring task. Fatigued participants tend to judge that morally elevating good deeds are less obligatory when the judgement focused on the self-sacrificial action compared to when it focused on the beneficial outcome; there was no effect for non-fatigued participants, as Experiment 2 shows. The result corroborates the idea that participants who have engaged in a cognitively tiring task judge that a good action is less obligatory because they construct a simple model of the events that does not explicitly link the self-sacrificial action to its beneficial outcome.

When individuals are fatigued by tiring laboratory tasks, they make different moral judgements and feel worse about their judgements, compared to individuals who are not fatigued. We suggest that cognitive fatigue affects moral judgements because people construct a simpler model of events when they are fatigued, one that does not explicitly represent the links between the action and the outcome. An alternative explanation is that fatigued participants were less motivated to try to think about the moral dilemmas. However, the fatigued participants tended to judge that reasoning about the moral dilemmas was more difficult than non-fatigued participants, and their metacognitive perception of difficulty suggests they did at least attempt to think about the dilemmas.

We propose that the moral fatigue effect is consistent with results that show that moral judgement is susceptible to similar influences that affect reasoning and decision making more generally. In particular, we suggest that given that cognitive fatigue affects general reasoning tasks, the demonstration in our experiments that cognitive fatigue also affects moral reasoning tasks may be difficult to reconcile with suggestions that moral judgement is a unique and separate domain-specific faculty (e.g., Hauser, 2006; Mikhail, 2007). Many factors that affect reasoning and decision making in general also affect moral judgement, such as framing effects (e.g., Parkinson & Byrne, 2017b; Sinnott-Armstrong, 2008), foreign language effects (Costa et al., 2014; Geipel, Hadjichristidis, & Surian, 2016), processing fluency effects (Laham et al., 2009), and reasons for actions (Rai & Holyoak, 2010; Ritov & Baron, 1999). Moreover, individual differences in abilities such as working memory capacity, as well as in general cognitive style, also influence moral judgements (e.g., Bartels, 2008; Bartels & Pizarro, 2011; Moore et al., 2008), as does the presentation of multiple alternatives simultaneously rather than sequentially (Paharia et al., 2009; see also Lombrozo, 2009). The results thus corroborate suggestions that reasoning about moral matters relies on the same cognitive processes as reasoning about non-moral matters (e.g., Białek & De Neys, 2017; Bucciarelli & Johnson-Laird, 2005; Gubbins & Byrne, 2014; Parkinson & Byrne, 2018;

Wiegmann & Osman, 2017), such as the construction of a model that causally links the action to the outcome (e.g., Crockett, 2013; Cushman, 2013; Lagnado et al., 2013). Overall, the experiments reported here indicate that people reason differently about moral problems after they have completed cognitively exhausting tasks.

References

Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: The "other-praising" emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology*, *4*, 105–127.

Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychological Science*, *23*, 861–868.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A*, *49*, 5–28.

Baddeley, A. (2007). *Working memory, thought, and action* (*vol. 45*). Oxford, UK: Oxford University Press.

Baron, J. (2017). Utilitarian vs. deontological reasoning. In J. F. Bonnefon & B. Trémolière (Eds.), *Moral inferences* (pp. 137–152). Abingdon, UK: Taylor & Francis.

Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, *108*, 381–417.

Bartels, D. M., & Medin, D. L. (2007). Are morally motivated decision makers insensitive to the consequences of their choices? *Psychological Science*, *18*, 24–28.

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*, 154–161.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*, 1–15.

Baumeister, R. F., & Vohs, K. D. (2016). Misguided effort with elusive implications. *Perspectives on Psychological Science*, *11*, 574–575.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*, 351–355.

Białek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, *12*, 148–167.

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, *352*, 1573–1576.

Bucciarelli, M., & Daniele, M. (2015). Reasoning in moral conflicts. *Thinking & Reasoning*, *21*, 265–294.

Bucciarelli, M., & Johnson-Laird, P. N. (2005). Naïve deontics: A theory of meaning, representation, and reasoning. *Cognitive Psychology*, *50*, 159–193.

Bucciarelli, M., Khemlani, S., & Johnson-Laird, P. N. (2008). The psychology of moral reasoning. *Judgment and Decision Making*, *3*, 121–139.

Capraro, V., & Cococcioni, G. (2016). Rethinking spontaneous giving: Extreme time pressure and ego-depletion favor self-regarding reactions. *Scientific Reports*, *6*, Article no. 27219.

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Frontiers in Psychology*, *5*, 823.

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

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS ONE*, *9*, e94842.

Cox, K. S. (2010). Elevation predicts domain-specific volunteerism 3 months later. *The Journal of Positive Psychology*, *5*, 333–341.

Crockett, M. J. (2013). Models of morality. *Trends in Cognitive Sciences*, *17*, 363–366.

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

Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: The aversion to harmful action. *Emotion*, *12*, 2–7.

Cushman, F., Knobe, J., & Sinnott-Armstrong, W. (2008). Moral appraisals affect doing/allowing judgments. *Cognition*, *108*, 281–289.

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*, 6–21.

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*, 1082–1089.

Damasio, A. R. (2000). A second chance for emotion. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 12–23). Oxford, UK: Oxford University Press.

Dang, J. (2016). Commentary: A multilab preregistered replication of the ego-depletion effect. *Frontiers in Psychology*, *7*, 1115.

Danzigera, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*, 6889–6892.

Freeman, D., Aquino, K., & McFerran, B. (2009). Overcoming beneficiary race as an impediment to charitable donations: Social dominance orientation, the experience of moral elevation, and donation behavior. *Personality and Social Psychology Bulletin*, *35*, 72–84.

Geipel, J., Hadjichristidis, C., & Surian, L. (2016). Foreign language affects the contribution of intentions and outcomes to moral judgment. *Cognition*, *154*, 34–39.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*, 509–517.

Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, *115*, 191–203.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*, 1029.

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, *23*, 206–215.

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*, 505–520.

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.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144–1154.

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.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108.

Gubbins, E., & Byrne, R. M. (2014). Dual processes of emotion and reason in judgments about moral dilemmas. *Thinking & Reasoning*, *20*, 245–268.

Gürçay, B., & Baron, B. (2017). Challenges for the sequential two-system model of moral judgement. *Thinking & Reasoning*, *23*, 49–80.

Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A., & Zwienenberg, M. (2016). A multi-lab pre-registered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*, 546–573.

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*, 495–525.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834.

Hauser, M. D. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York, NY: Ecco.

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology*, *43*, 497–504.

Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, *7*, 450–463.

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General*, *137*, 691–705.

Kahane, G., Everett, J. A., Earp, B. D., Farias, M., & Savulescu, J. (2015). "Utilitarian" judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193–209.

Kahane, G., & Shackel, N. (2010). Methodological issues in the neuroscience of moral judgement. *Mind & Language*, *25*, 561–582.

Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, *7*, 708–714.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 31–53). New York, NY: Holt, Rinehart and Winston.

Lagnado, D. A., Gerstenberg, T., & Zultan, R. I. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, *37*, 1036–1073.

Laham, S. M., Alter, A. L., & Goodwin, G. P. (2009). Easy on the mind, easy on the wrongdoer: Discrepantly fluent violations are deemed less morally wrong. *Cognition*, *112*, 462–466.

Lai, C. K., Haidt, J., & Nosek, B. A. (2014). Moral elevation reduces prejudice against gay men. *Cognition & Emotion*, *28*, 781–794.

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, *10*, 518–536.

Landy, J. F., & Royzman, E. B. (2018). The moral myopia model. In G. Pennycook (Ed.), *The new reflectionism in cognitive psychology: Why reason matters* (pp. 70–92). New York, NY: Psychology Press.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339–354.

Lombrozo, T. (2009). The role of moral commitments in moral judgment. *Cognitive Science*, *33*, 273–286.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science*, *341*, 976–980.

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, *55*, 102–111.

Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., & Ariely, D. (2009). Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology*, *45*, 594–597.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, *11*, 143–152.

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.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*, 247–259.

Muraven, M., Pogarsky, G., & Shmueli, D. (2006). Self-control depletion and the general theory of crime. *Journal of Quantitative Criminology*, *22*, 263–277.

Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition*, *100*, 530–542.

O' Hara, R.E., Sinnott-Armstrong, W., & Sinnott-Armstrong, N.A. (2010). Wording effects in moral judgements. *Judgment and Decision Making*, *5*, 547–554.

Paharia, N., Kassam, K. S., Greene, J. D., & Bazerman, M. H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes*, *109*, 134–141.

Parkinson, M., & Byrne, R. M. J. (2017a). Counterfactual and semi factual thoughts in moral judgments about failed attempts to harm. *Thinking & Reasoning*, *23*, 409–448.

Parkinson, M., & Byrne, R. M. J. (2017b). Moral judgments of risky choices: A moral echoing effect. *Judgment and Decision Making*, *12*, 236–252.

Parkinson, M., & Byrne, R. M. J. (2018). Judgments of moral responsibility and wrongness for intentional and accidental harm and purity violations. *Quarterly Journal of Experimental Psychology*, *71*, 779–789.

Patil, I. (2015). Trait psychopathy and utilitarian moral judgement: The mediating role of action aversion. *Journal of Cognitive Psychology*, *27*, 349–366.

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, *36*, 163–177.

Piaget, J. (1932). *The moral development of the child*. London, England: Kegan Paul.

Pizarro, D. A., & Salovey, P. (2002). Being and becoming a good person: The role of emotional intelligence in moral development and behavior. In J. Aronson (Eds.), *Improving academic achievement: Impact of psychological factors on education* (pp. 248–262). San Diego, CA: Academic Press.

Pohling, R., & Diessner, R. (2016). Moral elevation and moral beauty: A review of the empirical literature. *Review of General Psychology*, *20*, 412–425.

Rai, T. S., & Holyoak, K. J. (2010). Moral principles or consumer preferences? Alternative framings of the trolley problem. *Cognitive Science*, *34*, 311–321.

Ritov, I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes*, *79*, 79–94.

Royzman, E. B., Landy, J. F., & Leeman, R. F. (2015). Are thoughtful people more utilitarian? CRT as a unique predictor of moral minimalism in the dilemmatic context. *Cognitive Science*, *39*, 325–352.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*, 574–586.

Schmeichel, B. J. (2007). Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology: General*, *136*, 241–255.

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology*, *85*, 33–46.

Schnall, S., Roper, J., & Fessler, D. M. (2010). Elevation leads to altruistic behavior. *Psychological Science*, *21*, 315–320.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366.

Sinnott-Armstrong, W. (2008). Framing moral intuitions. In W. Sinnott-Armstrong (Eds.), *Moral psychology: The cognitive science of morality: Intuition and diversity* (Vol. 2) (pp. 47–76). Cambridge: The MIT Press.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661. doi:10.1126/science.283.5408.1657

Spears, D. E. (2010). *Economic decision-making in poverty depletes behavioral control*. Princeton, NJ: Center for Economic Policy Studies, Princeton University.

Tasso, A., Sarlo, M., & Lotto, L. (2017). Emotions associated with counterfactual comparisons drive decision-making in Footbridge-type moral dilemmas. *Motivation and Emotion*, *41*, 410–418.

Trémolière, B., De Neys, W., & Bonnefon, J. F. (2012). Mortality salience and morality: Thinking about death makes people less utilitarian. *Cognition*, *124*, 379–384.

Ugazio, G., Lamm, C., & Singer, T. (2012). The role of emotions for moral judgments depends on the type of emotion and moral scenario. *Emotion*, *12*, 579–590.

Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, *17*, 476–477.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). A dual-process specification of causal conditional reasoning. *Thinking & Reasoning*, *11*, 239–278.

Ward, A., & Mann, T. (2000). Don't mind if I do: Disinhibited eating under cognitive load. *Journal of Personality and Social Psychology*, *78*, 753–763.

Wiegmann, A., & Osman, M. (2017). Factors guiding moral judgment, reason, decision and action. *Experimental Psychology*, *64*, 65–67.

Wiegmann, A., & Waldmann, M. R. (2014). Transfer effects between moral dilemmas: A causal model theory. *Cognition*, *131*, 28–43.

Yilmaz, O., & Saribay, S. A. (2017). Activating analytic thinking enhances the value given to individualizing moral foundations. *Cognition*, *165*, 88–96.