



OPEN ACCESS

EDITED BY

Keith Markman,
Ohio University, United States

REVIEWED BY

Justin Landy,
Nova Southeastern University, United States
Piers D. L. Howe,
The University of Melbourne, Australia

*CORRESPONDENCE

Paul Rehren
✉ p.h.p.rehren@uu.nl

RECEIVED 20 February 2024

ACCEPTED 16 April 2024

PUBLISHED 02 May 2024

CITATION

Rehren P (2024) The effect of cognitive load, ego depletion, induction and time restriction on moral judgments about sacrificial dilemmas: a meta-analysis.
Front. Psychol. 15:1388966.
doi: 10.3389/fpsyg.2024.1388966

COPYRIGHT

© 2024 Rehren. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The effect of cognitive load, ego depletion, induction and time restriction on moral judgments about sacrificial dilemmas: a meta-analysis

Paul Rehren*

Ethics Institute, Department of Philosophy and Religious Studies, Utrecht University, Utrecht, Netherlands

Greene's influential dual-process model of moral cognition (mDPM) proposes that when people engage in Type 2 processing, they tend to make consequentialist moral judgments. One important source of empirical support for this claim comes from studies that ask participants to make moral judgments while experimentally manipulating Type 2 processing. This paper presents a meta-analysis of the published psychological literature on the effect of four standard cognitive-processing manipulations (cognitive load; ego depletion; induction; time restriction) on moral judgments about sacrificial moral dilemmas [$n = 44$; $k = 68$; total $N = 14,003$; $M(N) = 194.5$]. The overall pooled effect was in the direction predicted by the mDPM, but did not reach statistical significance. Restricting the dataset to effect sizes from (high-conflict) personal sacrificial dilemmas (a type of sacrificial dilemma that is often argued to be best suited for tests of the mDPM) also did not yield a significant pooled effect. The same was true for a meta-analysis of the subset of studies that allowed for analysis using the process dissociation approach [$n = 8$; $k = 12$; total $N = 2,577$; $M(N) = 214.8$]. I argue that these results undermine one important line of evidence for the mDPM and discuss a series of potential objections against this conclusion.

KEYWORDS

moral judgment, dual-process model, sacrificial dilemmas, cognitive-processing manipulations, meta-analysis

1 Introduction

Greene's *dual-process model of moral cognition* (mDPM¹; Greene et al., 2001; Greene, 2008, 2014) is one of the most well-known and influential models in moral psychology. Like other dual-process models, the mDPM makes the core assumption that "cognitive tasks evoke two forms of processing that contribute to observed behavior" (Evans and Stanovich, 2013, p. 225). These two forms of processing go by a variety of different names; here, I will refer to them as *Type 1* and *Type 2* processing. Ways to spell out the distinction between Type 1 and Type 2 processing abound (see e.g., Evans, 2008;

1 Short for moral dual-process model.

Evans and Stanovich, 2013). According to Greene et al. (2008, p. 40–1) and Greene (2014, p. 698–9), Type 1 processes are quick, effortless and unconscious, while Type 2 processes are slower, require effort and operate at the level of consciousness.

The central idea of the mDPM is that Type 1 and Type 2 processing tend to produce different kinds of moral judgment. When people engage in Type 1 processing, then they typically make *deontological* moral judgments—moral judgments that are “naturally justified in deontological terms (in terms of rights, duties, etc.) and that are more difficult to justify in consequentialist terms.”² In contrast, when people engage in Type 2 processing, then they typically make *consequentialist* moral judgments—moral judgments that are “naturally justified in consequentialist terms (i.e., by impartial cost-benefit reasoning) and that are more difficult to justify in deontological terms because they conflict with our sense of people’s rights, duties, and so on” (both, Greene, 2014, p. 699).

One way to illustrate the mDPM is with two classic *sacrificial moral dilemmas*. Sacrificial dilemmas are a type of moral dilemma; they involve situations where the agent can either do nothing or intervene (see e.g., Kahane et al., 2018, p. 132; Klenk, 2022, p. 593–594). If the agent does nothing, this will result in harm to or the death of one or more individuals. If the agent does intervene, she will save some or all of these individuals; however, the intervention will also cause harm, though less harm than would result from the agent doing nothing (e.g., a smaller number of individuals get killed). In one classic example (*Switch*), an out-of-control trolley is speeding toward five people. The only way to save them is to divert the trolley onto a separate track by hitting a switch. However, there is a sixth person on this other track who will then be run over and killed by the trolley if it is diverted (Foot, 1967). Another classic example (*Footbridge*) again features an out-of-control trolley; this time, however, the only way to save the five people is to push a sixth person off a footbridge and into the path of the trolley (Thomson, 1976). People tend to give the consequentialist response to *Switch* (that is, they endorse hitting the switch) but the deontological response to *Footbridge* (that is, they do not endorse pushing the person off the footbridge; e.g., Hauser et al., 2007; Awad et al., 2020).

To explain this finding, Greene et al. (2001, 2004) have argued that when most people consider *Footbridge*, this causes them to have a strong Type 1 response, one that only rarely gets overridden by subsequent Type 2 processing. Therefore, most people end up endorsing the deontological option in *Footbridge*. In contrast, considering *Switch* only causes a weak Type 1 response or no immediate Type 1 response at all, and so more people choose the consequentialist option in *Switch* than in *Footbridge*.

² Greene has also argued that deontological moral judgments (but not consequentialist moral judgments) are based on *emotional* processing. Note, however, that this claim is conceptually and empirically distinct from the claim that deontological moral judgments are typically the output of Type 1 processing (even though Greene does not always clearly distinguish the two claims; see Kahane, 2012, p. 522–3). In this paper, I am not concerned with the role of the emotions in deontological and consequentialist moral judgment.

One important source of evidence for the mDPM (e.g., Kahane, 2012, p. 524–526; Greene, 2014, p. 700–705; Guglielmo, 2015, p. 10–11) comes from studies that investigate the effect on moral judgments of experimental manipulations designed to either encourage participants to engage in Type 2 processing or to inhibit their ability to engage in Type 2 processing. Common methods to encourage Type 2 processing include direct instruction and time delays; common methods to inhibit Type 2 processing include cognitive load, time pressure and direction instruction (see Horstmann et al., 2010; Isler and Yilmaz, 2023). Most of these studies feature moral judgments about sacrificial moral dilemmas. The mDPM predicts that when participants are encouraged to engage in Type 2 processing, this will increase consequentialist responding, while inhibiting Type 2 processing will reduce consequentialist responding. However, previous overviews of the evidence have concluded that this body of studies only provides “inconsistent support” (Patil et al., 2021, p. 445) for the mDPM: Some studies have borne out the mDPM’s predictions, but others have not. To what extent this supports the mDPM overall, then—if it does indeed support it at all—is to date unknown.

In this paper, I aim to change this. To this end, I present a meta-analysis of published English-language psychological studies with adult participants investigating the effect of four standard cognitive-processing manipulations (cognitive load; ego depletion; induction; time restriction) on moral judgments about sacrificial moral dilemmas. The meta-analysis includes a total of 44 articles, reporting results from 68 individual studies [total $N = 14,003$; $M(N) = 194.5$]. Section 2 provides details about my inclusion criteria and the literature search. Section 3 then presents the meta-analysis, along with an analysis of the potential presence of publication bias and a series meta-regressions (manipulation type; experimental design; sample location; sample type; dilemma type; manipulation check; type of literature). To end, I discuss the implications of my results for the mDPM.

2 Method

2.1 Inclusion criteria

I searched for and included published English-language studies with adult participants investigating the effect of four standard cognitive-processing manipulations (cognitive load; ego depletion; induction; time restriction) on moral judgments about sacrificial moral dilemmas. I defined moral judgments as judgments about what one should or should not do in a situation, where this includes judgments about an action’s or a decision’s wrongness, permissibility, appropriateness and acceptability. This means that studies which measured moral behavior instead of moral judgment (e.g., offers in an economic game), asked participants to make a hypothetical choice (e.g., Would you push the button?), asked participants to indicate a preference (e.g., Which outcome would you prefer?) or asked participants to make a judgment about themselves, are beyond the scope of this review (see McDonald et al., 2021, p. 3). In addition, I excluded studies if they did not report original empirical data (e.g., review articles), only reported data from non-adult participants, had been published in a language other than English or had been retracted.

Various approaches have been tried to encourage and inhibit Type 2 processing (for overviews, see Horstmann et al., 2010; Isler and Yilmaz, 2023). Like two other recent meta-analyses (Rand, 2016; Kvarven et al., 2020), I focused on four standard cognitive-processing manipulations: cognitive load, ego depletion, induction and time restriction.

Cognitive load studies ask some participants to engage in a difficult cognitive task at the same time that they are making moral judgments. There are a variety of these tasks for researchers to choose from; common examples include tasks that tax participants' memory (e.g., memorize a complicated pattern of dots on a 4 × 4 grid) and tasks that require participants to pay attention to an additional stimulus (e.g., listen to a list of numbers and count how many prime numbers there are). The responses of participants under cognitive load are then compared to the responses of participants who engaged in an easier version of the load task (e.g., memorize a simple pattern of dots on a 3 × 3 grid) or no additional task, with the idea being that participants under cognitive load have fewer cognitive resources available for the moral judgment task and so will be inhibited from engaging in Type 2 processing while completing the task (Gilbert et al., 1988).

Ego depletion studies also aim to limit the amount of cognitive resources that participants have available to them while rendering moral judgments. They achieve this either by asking participants to complete a taxing cognitive task prior to the moral judgment task, or through mental or physical exertion (e.g., sleep deprivation; stress; hunger). Compared to participants who have completed a milder version of the depletion task or no task at all, depleted participants are supposed to be more mentally fatigued and so are less able or willing to engage in Type 2 processing (Hagger et al., 2010; but see Carter et al., 2019).

Induction studies encourage or discourage participants to engage in Type 2 processing while making moral judgments. One common way to achieve this is to give participants explicit instructions (e.g., think carefully and logically before making your judgment; think about how you would justify your judgment). Other induction studies have relied on unconscious primes. Many different primes have been tried, including memory recall (e.g., asking participants to think back to a time when reflection led to a favorable outcome), tasks like the Cognitive Reflection Test and the Berlin Numeracy Test (Cokely et al., 2012), and deliberative mindset primes.³

Finally, *time restriction* studies either limit the amount of time that participants have to make their moral judgment, or only allow participants to make their judgment after a time delay. Time-pressured and time-delayed responses are then compared with each other, or with the responses from participants whose response time was unconstrained. Time pressure makes it less likely that participants engage in Type 2 processing because they lack the time to do so. Conversely, time delays are thought to make it more likely

that participants will engage in Type 2 processing to make their judgment (Wright, 1974).

2.2 Literature search

The literature search was carried out in two steps (for details, see Figure 1). The first step was a database search; in the second step, I carried out backward and forward literature searches on all articles identified in the first step (plus some additional articles, see below). For the database search, I searched for relevant studies within *PubMed*, *PsycINFO* and *Web of Science*, using an intentionally broad combination of search terms like “reflection”, “deliberation”, “moral judgment” and “moral decision-making” (for details, see: osf.io/h2fcp/). This search yielded 3,451 unique records, which I then screened based on their titles. Finally, I assessed the full texts of the remaining records for eligibility. This left 24 peer reviewed articles to be included in the review.

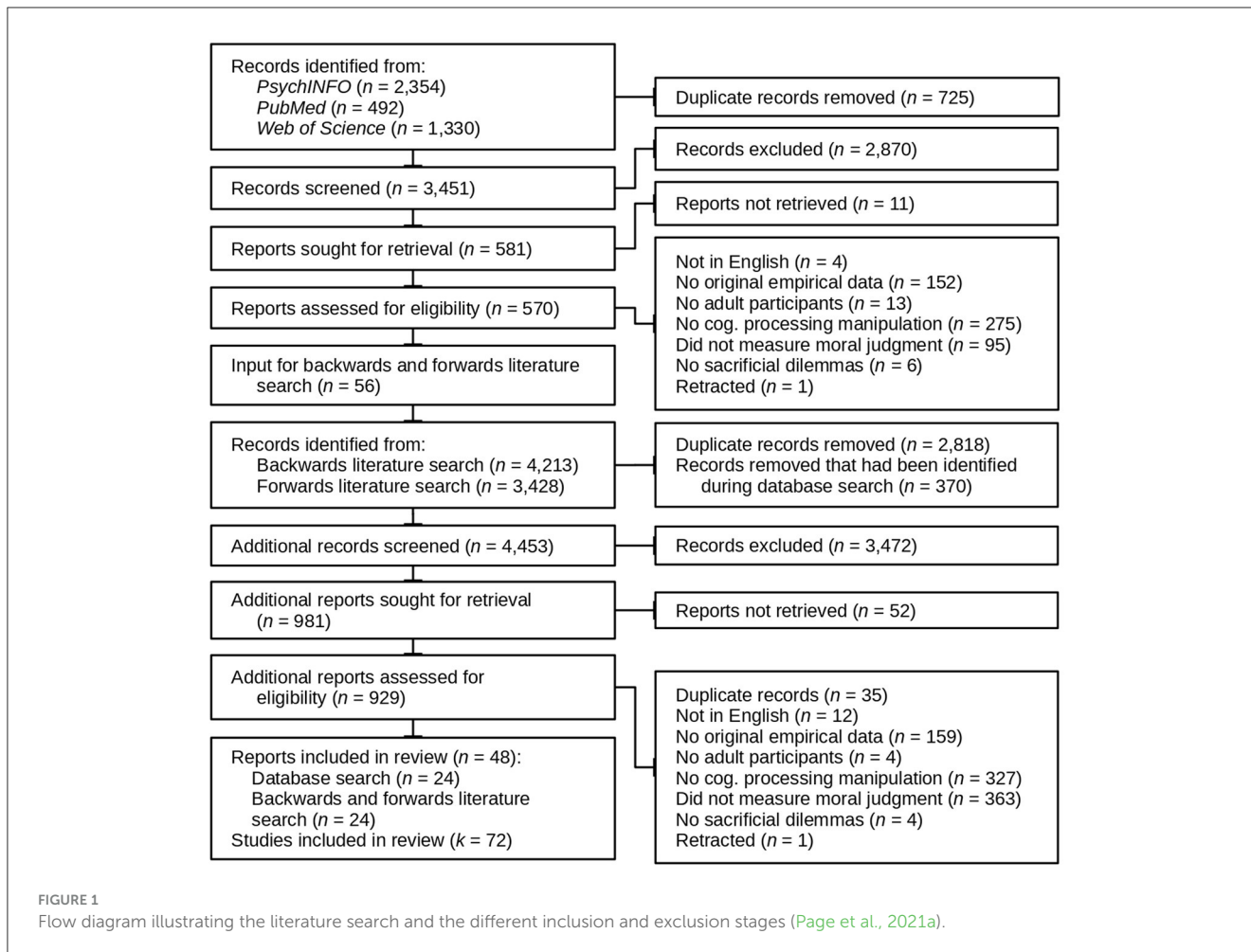
The second step consisted of a forward and backward literature search (pre-registered at: osf.io/gysuj/). These searches were carried out using Google Scholar, queried through Publish or Perish (Harzing, 2021). The input to this search were the 24 articles previously identified through database search, plus 26 additional articles that investigated the relationship between moral judgment and the tendency of participants to engage in intuitive and reflective processing (for an overview, see Patil et al., 2021) and six additional articles that investigated the relationship between moral judgment and cognitive processing with materials other than sacrificial dilemmas. These additional articles had also been identified in the initial database search. Forward and backward literature search resulted in an additional 4,418 unique records, which after screening and assessment for eligibility contributed an additional 24 articles.

All in all, I identified 48 articles that matched my inclusion criteria, which reported results from 72 individual studies.

2.3 Data collection

The majority of studies reported continuous outcomes. Therefore, I chose to use the bias-corrected standardized mean difference as the common effect size for the meta-analysis (Hedges' g_s for between-subject studies, Hedges' g_w for within-subject studies—I will drop the subscripts for the rest of the paper; Lakens, 2013). The remaining studies reported binary outcomes. Since these two sets of studies were otherwise very similar, I combined both in a single meta-analysis (see Borenstein et al., 2009, chap. 7). The meta-analysis included both between-subject and within-subject studies. To make these two sets of studies comparable, I converted all within-subject effect sizes to between-subject effect sizes (Morris and DeShon, 2002, Eq. 11). I coded all effect sizes so that positive values of g indicated results in line with the expectations of the mDPM: participants who were encouraged to engage in Type 2 processing made more consequentialist moral judgments than participants in a control condition or participants whose ability to engage in Type 2 processing was inhibited, while participants whose ability to engage in Type 2 processing was inhibited made

³ Recent reviews (Horstmann et al., 2010; Isler and Yilmaz, 2023) mention three other induction methods beyond explicit instruction and unconscious primes (enhancement of task relevance; monetary incentives; de-biasing training). However, I did not find any studies that employed any of these approaches and also matched the other inclusion criteria.



less consequentialist moral judgments than participants in a control condition.

Most studies reported sufficient statistical information to calculate g . For studies where this was not the case, I contacted the corresponding author(s) and requested the necessary statistical information. In the end, I was able to calculate g for all but four studies. None of the studies I was not able to include reported a significant effect of their cognitive-processing manipulation on participants' moral judgments.⁴ In the end, the meta-analysis included 44 articles, reporting results from 68 individual studies [total $N = 14,003$; $M(N) = 194.5$]. Table 1 shows a detailed overview of these studies.

3 Results

3.1 Point estimate

All analyses were carried out in R (R Core Team, 2023).⁵ To calculate a point estimate and 95% confidence interval for the

combined effect of cognitive load, ego depletion, induction and time restrictions on moral judgments about sacrificial dilemmas, I used a three-level random-effects model, with effect sizes nested within studies. This was done to account for the fact that some studies contributed multiple effect sizes (see Harrer et al., 2021, chap. 10). While there was a small (Cohen, 1988) positive pooled effect, it did not reach statistical significance: $g = 0.06$, 95% CI = (0.00, 0.12), $p = 0.057$. Figure 2 illustrates the results of the meta-analysis in a forest plot.⁶

There was evidence of considerable heterogeneity: $Q(108) = 433.17$; $p < 0.001$. $I^2_{\text{Level 3}} = 2.2\%$ of the total variation can be attributed to heterogeneity between studies, while $I^2_{\text{Level 2}} = 80.9\%$ can be attributed to heterogeneity within studies.

bookdown (Xie, 2023); readr (Wickham et al., 2023b), and rmarkdown (Allaire et al., 2023).

6 Three studies (Killgore et al., 2007; Cummins and Cummins, 2012; Paxton et al., 2014) reported at least one large effect size ($g > 0.8$; Cohen, 1988; recall that some studies reported more than one relevant effect size) in line with the predictions of the mDPM (see Figure 4). There were no obvious commonalities between these studies (for example, in terms of any of the attributes listed in Table 1).

TABLE 1 Detailed information about all studies included in the meta-analysis; “—” signifies that a piece of information was not reported in the article in question.

Study	Design	N	Type 2 processing manipulation				Sample		Manipulation check		Personal (%)	PD	Published
			Type	Encouraged	Control	Inhibited	Location	Type	Type	Status			
Bago and De Neys (2019), Study 4	Within	101	Induction; load; time	Instructions to consider the scenario carefully and to take one's time		Dot memorization task; time pressure	Europe	Online		None	50.0	Yes	Yes
Bialek and De Neys (2017), Study 2	Between	141	Load		Control	Dot memorization task	North America	Online		None	0.0	Yes	Yes
Bialek and De Neys (2017), Study 3	Between	151	Load		Control	Dot memorization task	North America	Online		None	0.0	Yes	Yes
Cellini et al. (2017)	Between	37	Depletion		Control	Sleep deprivation	Europe	Students	Compared self-reported sleepiness and fatigue between conditions	Success	50.0	No	Yes
Chaturapanich and Chaiwutikornwanich (2015), Study 1	Between	360	Induction; time	Instructions to think carefully		Time pressure	Other	Students		None	0.0	No	Yes
Chaturapanich and Chaiwutikornwanich (2015), Study 2	Between	360	Induction; time	Instructions to think carefully		Time pressure	Other	Students		None	0.0	No	Yes
Conway and Gawronski (2013), Study 2	Between	56	Load		Control	String memorization task	North America	Students		None	50.0	Yes	Yes
Cova et al. (2018), 2008—most cited	Within	85	Load		Control	Digit attention task	Europe	Other		None	100.0	No	Yes
Cova et al. (2018), 2012—most cited	Between	298	Induction	Cognitive Reflection Test	Control		Europe	Online		None	100.0	No	Yes
Cummins and Cummins (2012), Study 1	Between	182	Time		Control	Time pressure	North America	Students	Compared response times between conditions	Success	64.7	No	Yes

(Continued)

TABLE 1 (Continued)

Study	Design	N	Type 2 processing manipulation			Sample		Manipulation check		Personal (%)	PD	Published	
			Type	Encouraged	Control	Inhibited	Location	Type	Type				Status
Cummins and Cummins (2012), Study 2	Within	65	Time		Control	Time pressure	North America	Students	Compared response times between conditions	Success	64.7	No	Yes
Doerflinger and Gollwitzer (2020), Study 1	Between	204	Induction	Deliberative mindset prime	Control		—	Online		None	100.0	No	Yes
Gamez-Djokic and Molden (2016), Study 7	Within	51	Load		Control	String memorization task	North America	Students		None	50.0	No	Yes
Gamez-Djokic and Molden (2016), Study 8	Within	144	Induction	Instructions to think carefully and logically and to take one's time	Control		—	Online		None	50.0	No	Yes
Gamez-Djokic and Molden (2016), Study 9	Within	143	Induction	Instructions to think carefully and logically and to take one's time	Control		—	Online		None	50.0	No	Yes
Gawronski et al. (2017), Study 2a	Between	194	Load		Control	String memorization task	—	Online	Compared memorization accuracy between conditions	Success	66.7	Yes	Yes
Gawronski et al. (2017), Study 2b	Between	194	Load		Control	String memorization task	—	Online	Compared memorization accuracy between conditions	Success	66.7	Yes	Yes
Greene et al. (2008)	Between	82	Load		Control	Digit attention task	North America	Students		None	100.0	No	Yes
Gürçay and Baron (2017), Study 1	Within	64	Induction	Instructions to make reflective judgments		Instructions to make intuitive judgments	—	Online	Compared response times between conditions	Success	50.0	No	Yes
Gürçay and Baron (2017), Study 2	Between	55	Induction	Instructions to make reflective judgments		Instructions to make intuitive judgments	—	Online	Compared response times between conditions	Success	50.0	No	Yes

(Continued)

TABLE 1 (Continued)

Study	Design	N	Type 2 processing manipulation				Sample		Manipulation check		Personal (%)	PD	Published
			Type	Encouraged	Control	Inhibited	Location	Type	Type	Status			
Gürçay and Baron (2017), Study 3	Between	194	Induction; time	Instructions to make reflective judgments; time delay		Instructions to make intuitive judgments; time pressure	—	Online	Compared response times between conditions	Success	47.8	No	Yes
Gürçay and Baron (2017), Study 3	Within	85	Induction; time	Instructions to make reflective judgments; time delay		Instructions to make intuitive judgments; time pressure	—	Online	Compared response times between conditions	Failed	47.8	No	Yes
Ham and van den Bos (2010)	Between	100.0	Induction; time	Instructions to deliberate; time delay	Control		Europe	Students		None	100.0	No	Yes
Hashimoto et al. (2022)	Within	119	Time		Control	Time pressure	Other	Students		None	0.0	No	Yes
Houston (2010), Study 1	Between	189	Load		Control	String memorization task	North America	Other		None	100.0	No	No
Houston (2010), Study 2	Between	292	Load		Control	String memorization task	North America	Other		None	100.0	No	No
Houston (2010), Study 3	Between	106	Load		Control	String memorization task	North America	Other		None	100.0	No	No
Houston (2010), Study 4	Between	99	Time		Control	Time pressure	North America	Other		None	100.0	No	No
Killgore et al. (2007)	Within	26	Depletion		Control	Sleep deprivation	North America	Other		None	50.0	No	Yes
Kroneisen and Steghaus (2021), Study 1	Between	199	Time	Time delay		Time pressure	Europe	Other	Compared response times between conditions	Success	75.0	Yes	Yes
Kroneisen and Steghaus (2021), Study 2	Between	168	Induction	Instructions to take their time and deliberate		Instructions to answer quickly and intuitively	Europe	Other	Compared response times between conditions	Failed	66.7	Yes	Yes

(Continued)

TABLE 1 (Continued)

Study	Design	N	Type 2 processing manipulation			Sample		Manipulation check		Personal (%)	PD	Published	
			Type	Encouraged	Control	Inhibited	Location	Type	Type				Status
Körner and Volk (2014), Study 1	Between	112	Time	Time delay		Time pressure	—	Online		None	100.0	No	Yes
Li et al. (2018), Study 2	Between	43	Induction; load	Instructions to deliberate		Letter memorization task; instructions to make one's decision based on feeling	Other	Students	Compared memorization accuracy between conditions	Success	66.7	Yes	Yes
Li et al. (2021)	Between	78	Depletion		Control	Trier Social Stress Test	Other	Students	Compared self-reported stress and heart rate between conditions	Success	66.7	Yes	Yes
Liu and Liao (2021), Study 3	Between	180	Induction; load	Instructions to think about the reasons for one's decision		Letter memorization task	Other	Students		None	100.0	No	Yes
Lyrantzis (2017), Study 1c	Between	104	Induction	Cognitive Reflection Test	Control		North America	Online		None	100.0	No	No
McPhetres et al. (2018), Study 1	Between	220	Time		Control	Time pressure	—	Other		None	50.0	Yes	Yes
McPhetres et al. (2018), Study 2	Between	186	Load		Control	String memorization task	North America	Students		None	50.0	Yes	Yes
McPhetres et al. (2018), Study 3	Between	801	Load		Control	String memorization task	North America	Online		None	50.0	Yes	Yes
Paxton et al. (2012), Study 1a	Between	92	Induction	Cognitive Reflection Test	Control		—	Online		None	100.0	No	Yes
Paxton et al. (2012), Study 1b	Between	72	Induction	Cognitive Reflection Test	Control		North America	Students		None	100.0	No	Yes
Paxton et al. (2014)	Between	17	Induction	Cognitive Reflection Test	Control		—	Online		None	100.0	No	Yes

(Continued)

TABLE 1 (Continued)

Study	Design	N	Type 2 processing manipulation				Sample		Manipulation check		Personal (%)	PD	Published
			Type	Encouraged	Control	Inhibited	Location	Type	Type	Status			
Reid (2018)	Between	166	Load		Control	String memorization task	Other	Other	Compared self-reported task difficulty between conditions	Success	62.5	No	No
Rosas and Aguilar-Pardo (2020), Study 1	Between	344	Induction; time	Instructions to consider the scenario carefully and to take one's time; time delay		Instructions to think fast; time pressure	—	Online		None	100.0	No	Yes
Rosas and Aguilar-Pardo (2020), Study 2	Between	709	Induction; time	Instructions to consider the scenario carefully and to take one's time; time delay		Instructions to think fast; time pressure	—	Online		None	100.0	No	Yes
Rosas and Aguilar-Pardo (2020), Study 3	Between	657	Induction; time	Instructions to consider the scenario carefully and to take one's time; time delay		Instructions to think fast; time pressure	—	Online		None	100.0	No	Yes
Rosas and Aguilar-Pardo (2020), Study 4	Between	554	Induction; time	Instructions to consider the scenario carefully and to take one's time; time delay		Instructions to think fast; time pressure	—	Online		None	100.0	No	Yes
Rosas et al. (2019), Study 3	Between	284	Load		Control	Dot memorization task	Other	Online	Compared response times between conditions	Success	100.0	No	Yes
Schwitzgebel and Cushman (2015)	Between	2686	Induction; time	Instructions to consider the scenario carefully and to take one's time; time delay	Control		North America	Other	Compared response times between conditions	Success	66.6	No	Yes
Simpson (2021)	Between	182	Induction		Control	Cognitive Reflection Test	North America	Students		None	100.0	No	No
Spears et al. (2021), Study 1a—BNT	Between	83	Induction	Berlin Numeracy Test with feedback	Control		Europe	Students	Compared rates of correct answers on the Berlin Numeracy Test before and after feedback	Success	100.0	No	Yes
Spears et al. (2021), Study 1a—CRT	Between	138	Induction	Cognitive Reflection Test with feedback	Control		Europe	Students	Compared rates of correct answers on the Cognitive Reflection Test before and after feedback	Success	100.0	No	Yes

(Continued)

TABLE 1 (Continued)

Study	Design	N	Type 2 processing manipulation				Sample		Manipulation check		Personal (%)	PD	Published
			Type	Encouraged	Control	Inhibited	Location	Type	Type	Status			
Spears et al. (2021), Study 1b—BNT	Between	99	Induction	Berlin Numeracy Test with feedback	Control		Europe	Online	Compared rates of correct answers on the Berlin Numeracy Test before and after feedback	Success	100.0	No	Yes
Spears et al. (2021), Study 1b—CRT	Between	86	Induction	Cognitive Reflection Test with feedback	Control		Europe	Online	Compared rates of correct answers on the Cognitive Reflection Test before and after feedback	Success	100.0	No	Yes
Starcke et al. (2012)	Between	50	Depletion		Control	Stress induction task	Europe	Students	Compared heart rate between conditions	Success	50.0	No	Yes
Suter and Hertwig (2011), Study 1	Between	67	Time	Time delay		Time pressure	Europe	Students	Compared response times between conditions	Success	50.0	No	Yes
Suter and Hertwig (2011), Study 2	Between	80	Induction	Instructions to take one's time to deliberate		Instructions to answer quickly and intuitively	Europe	Students	Compared response times between conditions	Success	50.0	No	Yes
Swann et al. (2014), Study 5	Between	436	Time		Control	Time pressure	Europe	Students		None	0.0	No	Yes
Timmons and Byrne (2019), Study 1	Between	196	Depletion		Control	Online depletion task	Other	Online	Compared self-reported task difficulty between conditions	Success	50.0	No	Yes
Timmons and Byrne (2019), Study 2	Between	187	Depletion		Control	Online depletion task	Other	Online	Compared self-reported task difficulty between conditions	Success	50.0	No	Yes
Tinghög et al. (2016), Study 1	Between	1100	Time	Time delay		Time pressure	Other	Other	Compared performance on Jellybean task between conditions	Success	50.0	No	Yes
Tinghög et al. (2016), Study 2	Between	311	Load		Control	Digit memorization task/attention depletion task	Europe	Students	Compared self-reports of how much energy the load task took between conditions	Success	50.0	No	Yes
Trémolière and Bonnefon (2014), Study 1	Between	213	Load		Control	Dot memorization task	Europe	Online		None	100.0	No	Yes

(Continued)

TABLE 1 (Continued)

Study	Design	N	Type 2 processing manipulation				Sample		Manipulation check		Personal (%)	PD	Published
			Type	Encouraged	Control	Inhibited	Location	Type	Type	Status			
Trémolière and Bonnefon (2014), Study 2	Between	123	Time		Control	Time pressure	Europe	Online		None	100.0	No	Yes
Trémolière and Bonnefon (2014), Study 3	Between	204	Time		Control	Time pressure	Europe	Online		None	75.0	No	Yes
Trémolière et al. (2012), Study 2	Between	110	Load		Control	Dot memorization task	Europe	Students		None	100.0	No	Yes
Van't Veer and Slegers (2019)	Between	185	Depletion		Control	Anticipatory stress task	Europe	Students	Compared self-reported stress between conditions	Success	0.0	No	Yes
Vega et al. (2021), Study 1	Between	59	Induction		Control	Instructions to give the first answer that comes to mind	—	Online	Compared response times between conditions	Success	100.0	No	Yes
Vega et al. (2021), Study 1	Within	30	Induction	Instructions to consider the scenario carefully and to take one's time		Instructions to give the first answer that comes to mind	—	Online	Compared response times between conditions	Success	100.0	No	Yes
Vega et al. (2021), Study 2	Within	119	Induction	Instructions to consider the scenario carefully and to take one's time		Instructions to give the first answer that comes to mind	—	Online	Compared response times between conditions	Success	50.0	No	Yes
Weippert et al. (2018)	Between	32	Depletion		Control	High intensity ergometer cycling session	Europe	Other	Compared cortisol levels and self-reported effort, pain, fatigue, mood, wakefulness and arousal between conditions	Success	50.0	No	Yes
Youssef et al. (2012)	Between	65	Depletion		Control	Trier Social Stress Test	Other	Students	Compared cortisol levels between conditions	Success	50.0	No	Yes
Zhao et al. (2016)	Between	159	Depletion		Control	Anxiety induction task	North America	Students	Compared self-reported anxiety between conditions	Success	50.0	No	Yes
Zheng (2018), Study 6	Between	137	Induction	Instructions to consider the scenario carefully and to take one's time	Control		Europe	Students		None	100.0	No	No

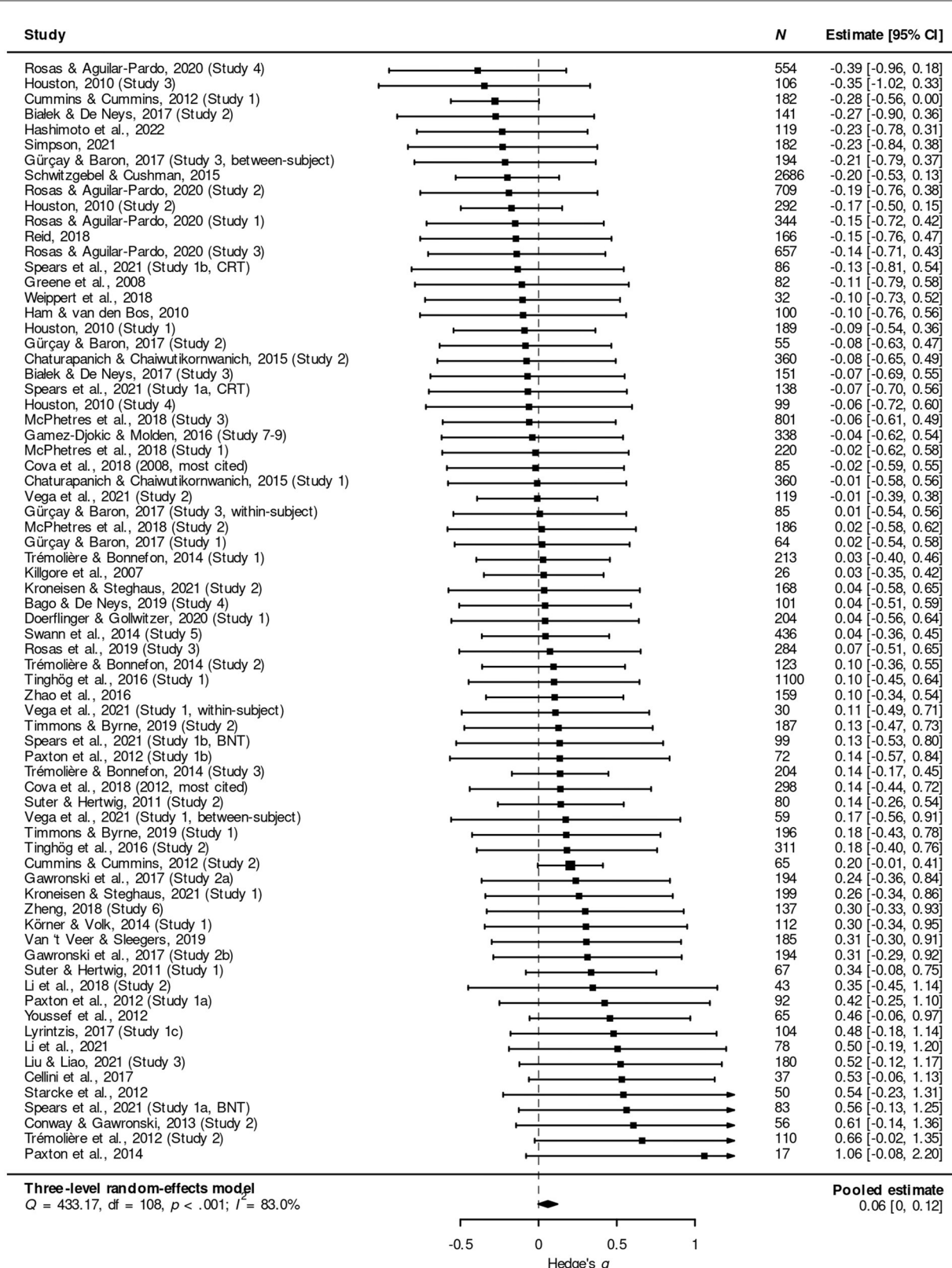


FIGURE 2 Forest plot showing the meta-analytic effect of four standard cognitive-processing manipulations on moral judgments about acts or decisions in sacrificial moral dilemmas. Positive values of g indicate results in line with the mDPM.

3.2 PD studies

A recent methodological critique of using sacrificial dilemmas to test the mDPM points out that in sacrificial moral dilemmas, participants can only choose between two options (intervene or do nothing). Therefore, studies that use these scenarios to test the mDPM need to assume that consequentialist and deontological responses are inversely related—when people endorse the consequentialist option in a sacrificial dilemma, they also reject (or would reject) the deontological option (and vice versa). However, deontology and consequentialism (the moral frameworks underlying these responses) are not related in this way, but are conceptually distinct. This is an issue for research on the mDPM, because even if it turned out that increased Type 2 processing results in more consequentialist responding and decreased Type 2 processing results in more deontological responding, this need not be evidence of two distinct types of cognitive processes, one consequentialist, one deontological. Instead, the same results are also consistent with single process accounts (e.g., Schein and Gray, 2018).

To do better, Conway and Gawronski (2013) have proposed to use *congruent sacrificial dilemmas* in addition to traditional sacrificial dilemmas. Like traditional dilemmas, congruent dilemmas pit a harmful action against a harmful inaction. However, in a congruent dilemma, the harm that would be caused if the agent intervened is greater than the harm that would result if the agent did nothing. The advantage of including both traditional and congruent dilemmas in a study is that its results can be analyzed using the *process dissociation* (PD) procedure (Jacoby et al., 1987). This procedure allows researchers to “independently quantify the strength of deontological and utilitarian inclinations” (Conway and Gawronski, 2013, p. 219). PD analysis outputs two quantities, U and D , where U measures the extent to which an individuals’ moral judgments (about sacrificial dilemmas) are based on consequentialism, while D does the same for deontology. The DPM predicts that encouraging participants to engage in Type 2 processing will lead to higher values of U , while inhibiting Type 2 processing will lead to lower values of U . Since all PD studies also include traditional sacrificial dilemmas, my literature search should also have identified all published PD studies (within the parameters described in Figure 1). Nine articles, in addition to traditional sacrificial dilemmas, also used the PD analysis approach ($k = 14$). Here, I present meta-analytic results for 8 of these articles, reporting results from 12 individual studies [total $N = 2,577$; $M(N) = 214.8$].⁷

Again, I used Hedge’s g , since the parameter U is continuous. g was once more coded so that positive values indicate results in line with mDPM: higher values of U for participants who were encouraged to engage in Type 2 processing compared to participants in a control condition or participants whose ability to engage in Type 2 processing was inhibited; lower values of U for participants whose ability to engage in Type 2 processing was inhibited compared to participants in a control condition. The pooled effect is $g = 0.11$, 95% CI = $(-0.01, 0.23)$, $p = 0.079$.

⁷ I was not able to include two studies (Biatek and De Neys, 2017, Study 2 & 3). Neither study reports finding a significant effect.

This effect is in the expected direction (from the perspective of the mDPM); however, it did not reach the conventional level of statistical significance ($p < 0.05$). Figure 3 illustrates the results in a forest plot.

3.3 Publication bias

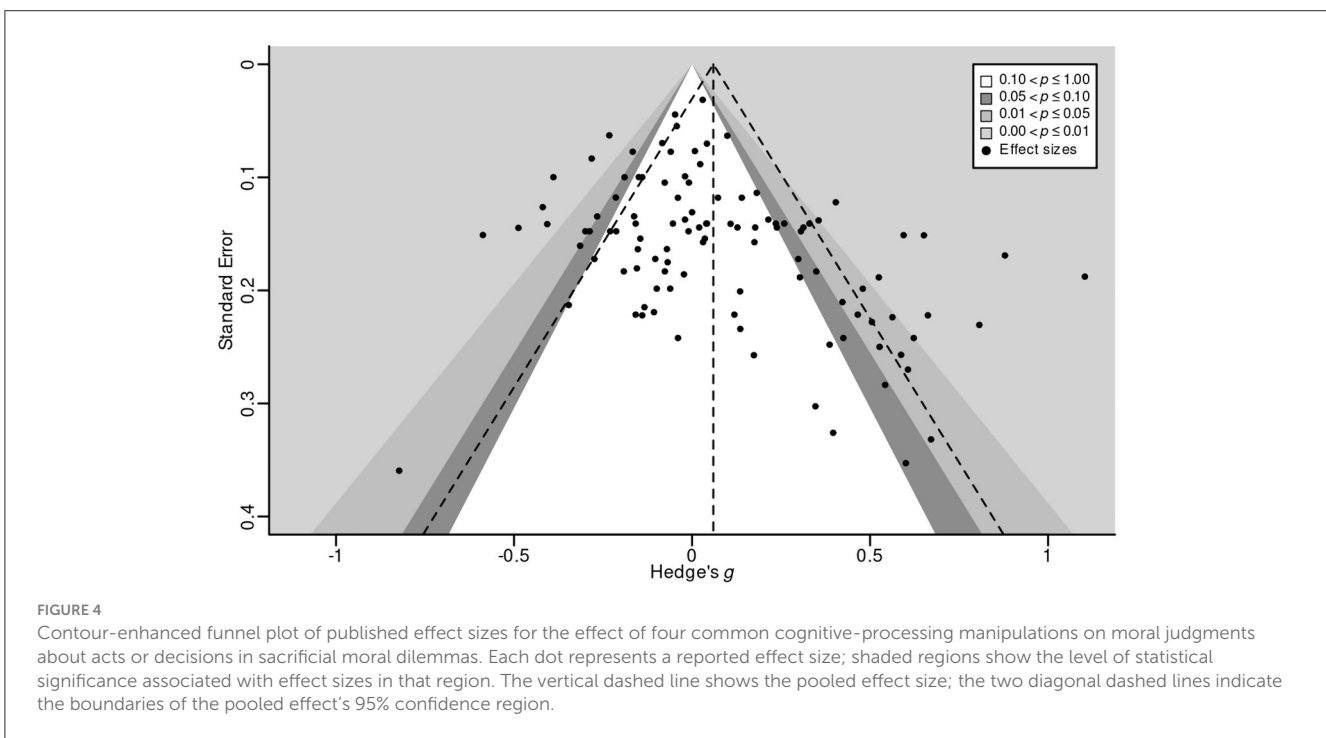
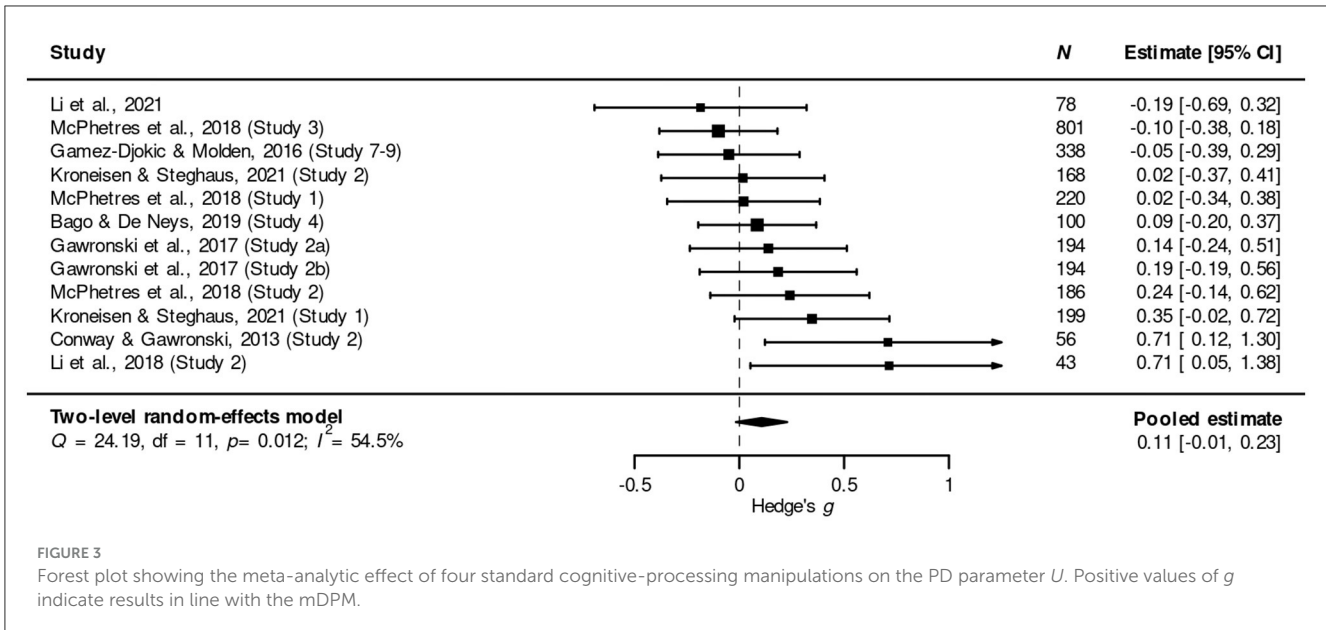
Publication bias happens when the probability that a study gets published are affected by its results. Two common reasons for publication bias are the greater tendency of significant results (compared to non-significant results) to see the light of day and the greater tendency of results that support the initial hypothesis to get published (Fanelli, 2012; Ferguson and Brannick, 2012; Franco et al., 2014; Kühberger et al., 2014). Publication bias can make it difficult to draw valid meta-analytic conclusions; in particular, its presence often results in inflated overall effect size estimates, since studies that find weaker or conflicting evidence will tend to be missing from the literature (see Thornton and Lee, 2000; Rothstein et al., 2005).

One common method to attempt to detect publication bias uses *funnel plots* (Peters et al., 2008). In a funnel plot, effect size (x -axis) is plotted against the inverse of the standard error of the effect size (y -axis). If there is no publication bias, then this plot should look roughly like an upside-down funnel, since effect sizes with low standard errors (i.e., high precision) are expected to cluster near the pooled effect size, while effect sizes that are associated with higher standard errors (i.e., lower precision) will be more widely dispersed.

Figure 4 shows a contour-enhanced funnel plot for the current meta-analysis (excluding gray literature). Visual inspection suggests substantial asymmetry. More specifically, there are almost no effect sizes in the bottom-left area of the plot, meaning that there are few published studies with high standard errors that have reported negative effects. Supporting this impression, Egger’s regression test (Egger et al., 1997) was significant, $\hat{\beta}_0 = -0.17$, $t = 3.22$, $p = 0.001$. Both results strongly suggest (but do not prove; see Page et al., 2021b) the presence of small- N publication bias, in the sense that studies with high standard errors likely had a lower probability of getting published if their results contradicted the expectation of the mDPM.

One standard way to estimate the “true” effect size (that is, the effect size in the absence of publication bias) is Duval and Tweedie (2000)’s *trim-and-fill* method. The idea behind this method is to impute the studies that are assumed to be missing from the data set until the funnel plot is symmetric. The pooled effect size of this extended data set then represents the estimate corrected for small-study publication bias. When applied to the current set of studies (number of imputed studies = 17), the estimated effect of cognitive-processing manipulation on moral judgments of sacrificial dilemmas becomes negative, but remains statistically indistinguishable from zero, $g = -0.02$, 95% CI = $(-0.09, 0.05)$, $p = 0.488$.

In addition, I also ran *PET-PEESE*, a more recent method to correct meta-analytic estimates for small-study publication bias (Stanley and Doucouliagos, 2014). The PET-PEESE-corrected estimate is significant, $g = -0.11$, 95% CI = $(-0.20, -0.02)$, $p = 0.007$, but in the opposite direction from what the mDPM predicts.



PET-PEESE is known to have a tendency to over-correct for bias, however, meaning that it may be more appropriate to interpret this result as suggesting the lack of a pooled effect instead of a negative pooled effect (Carter et al., 2019).

3.4 Meta-regression

Next, I examined the role of a series of potential moderators: type of cognitive-processing manipulation (Horstmann et al., 2010; Isler and Yilmaz, 2023); experimental design (between-subject vs.

within-subject; Morris and DeShon, 2002); geographic location of the sample (Henrich et al., 2010; Medin et al., 2017); participant pool (for example, students or online workers; Peterson, 2001; Stewart et al., 2017); type of sacrificial dilemma (impersonal or personal); the presence and status of a manipulation check; and type of literature (peer reviewed vs. gray). To this end, I used mixed-effects meta-regression, entering each potential moderator as a predictor into a three-level mixed-effects model (first level: individual effect sizes; second level: individual studies; third level: meta-analytic aggregate; see Harrer et al., 2021, sect. 10.1). Figure 5 illustrates the results in a forest plot.

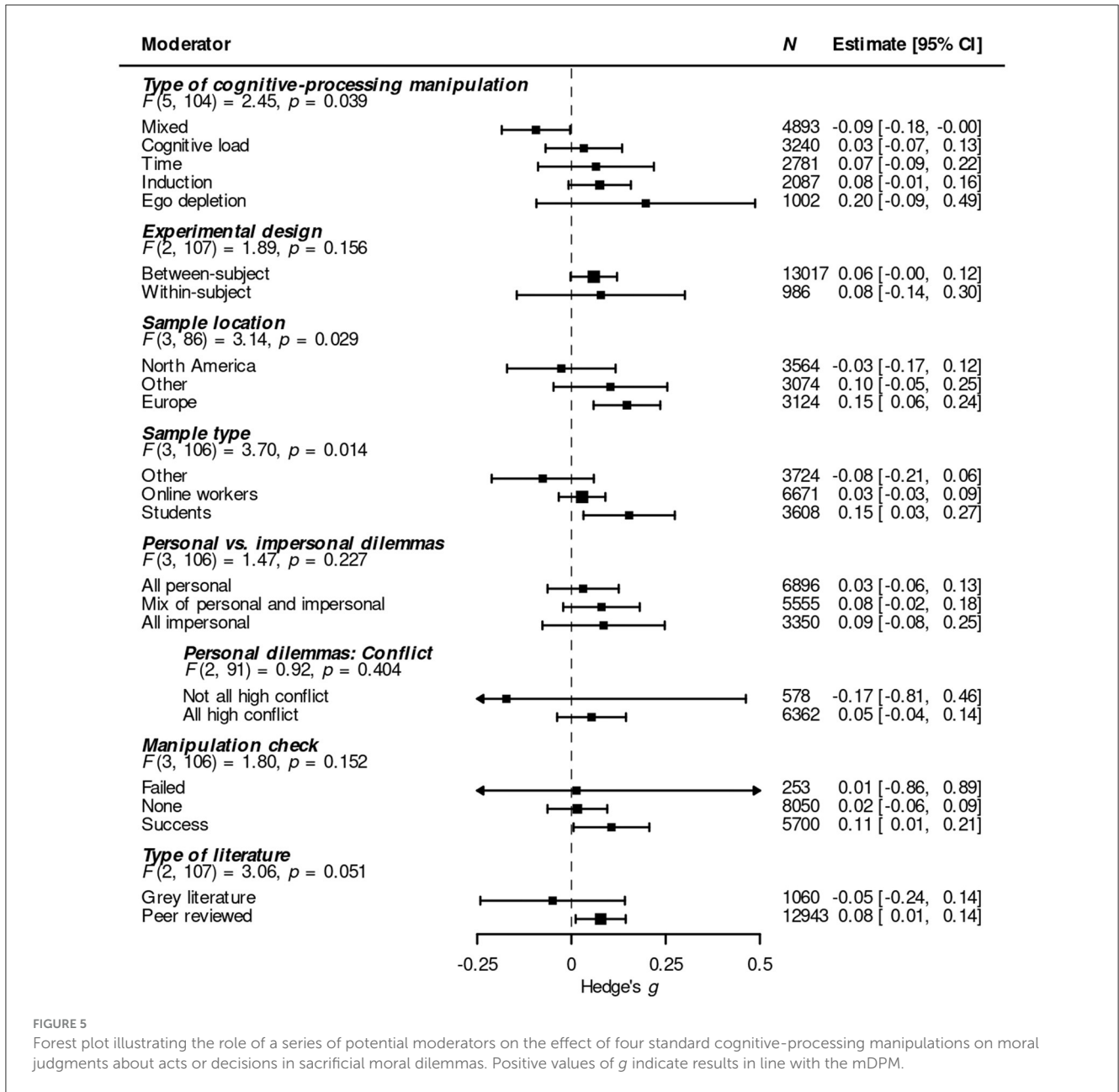


FIGURE 5 Forest plot illustrating the role of a series of potential moderators on the effect of four standard cognitive-processing manipulations on moral judgments about acts or decisions in sacrificial moral dilemmas. Positive values of *g* indicate results in line with the mDPM.

3.4.1 Type of cognitive-processing manipulation

Cognitive-processing manipulations likely differ in terms of their effectiveness (Horstmann et al., 2010; Isler and Yilmaz, 2023). For example, instead of inhibiting Type 2 processing, time pressure may in many cases reduce comprehension, making this manipulation less effective than other methods. Conversely, some induction manipulations (e.g., direct instructions) may be associated with considerable experimenter demand and so be more effective than other approaches.

Thirty-three effect sizes came from studies that used time restriction, followed by induction ($k = 22$), ego depletion ($k = 17$) and cognitive load ($k = 21$). Another 16 effect sizes had been found using a mix of more than one of the other four types (for details, see Table 1). Type of cognitive-processing manipulation had a significant overall effect on effect size: $F(5,104) = 2.45, p =$

0.039. The pooled effect of ego depletion studies was larger than that of the other manipulation types except for time restriction ($F_s \geq 4.03, p_s \leq 0.026$). Moreover, the pooled effect of mixed studies was smaller than that of studies using ego depletion and induction ($F_s \geq 4.39, p_s \leq 0.02$). It was also the only pooled effect statistically different from zero, though in the opposite direction of what the mDPM predicts: $g = -0.09, 95\% \text{ CI} = (-0.18, 0.00), p = 0.043$. None of the other pairwise differences reached statistical significance.

3.4.2 Experimental design

Most effect sizes came from studies that used a between-subject design ($k = 86$; within-subject: $k = 23$). The effect of experimental design was not significant, $F(2,107) = 1.89, p = 0.156$.

3.4.3 Sample location

As is common in social science research (Arnett, 2008; Rad et al., 2018; Thalmayer et al., 2021), a large majority of effect sizes came from samples of North Americans ($k = 41$) or Europeans ($k = 35$).⁸ In contrast, samples from other continents only contributed 13 effect sizes. Sample location had a significant on effect size, $F_{(3,86)} = 3.14$, $p = 0.029$. The pooled effect of samples from Europe was significantly larger than the pooled effect of samples from Other or North America ($F_s \geq 3.37$; $p_s \leq 0.04$) and was the only pooled effect that reached statistical significance, $g = 0.15$, 95% CI = (0.06, 0.24), $p = 0.002$. No treatment of the mDPM that I am aware of discusses the potential role of sample location as a mediator, let alone predicts the specific differences just described.

3.4.4 Sample type

Most effect sizes came from samples of online workers (from platforms like MTurk or Prolific; $k = 37$) or university students ($k = 48$; again, this is common for social science research: Arnett, 2008; Stewart et al., 2017; Thalmayer et al., 2021). Only 24 effect sizes were not from samples entirely made up of students or online workers. Type of subject pool had a significant overall effect on effect size, $F_{(3,106)} = 3.70$, $p = 0.014$. More specifically, the pooled effect in student samples was larger than the pooled effect in samples from the other two types of subject pool ($F_s \geq 4.18$, $p_s \leq 0.019$). The pooled effect for student samples was the only pooled effect that reached statistical significance [$g = 0.15$, 95% CI = (0.03, 0.27), $p = 0.014$]. No treatment of the mDPM that I am aware of discusses the potential role of sample type as a mediator, let alone predicts the specific differences found here.

3.4.5 Personal vs. impersonal dilemmas

Since the mDPM was first introduced, there has been considerable disagreement regarding how best to use sacrificial moral dilemmas to test the model.⁹ Greene et al. (2001) distinguished between two types of sacrificial dilemmas, *personal* and *impersonal* dilemmas. In a personal dilemma, the agent's choice must involve serious bodily harm to one or more particular individuals, where this harm is not the result of deflecting an existing threat. Footbridge is a paradigmatic example of a personal sacrificial dilemma. Impersonal dilemmas fail to meet at least one of these conditions; one paradigmatic example is Switch. Greene et al. (2008, p. 1146) then recommended that only personal dilemmas be used to test the mDPM.

I assigned each effect size to one of three categories based on the proportion of personal dilemmas used in the study it originated

⁸ Sample location was not reported for 20 effect sizes; these effect sizes were excluded from this analysis (for details, see Table 1).

⁹ Some researchers have argued that sacrificial dilemmas are very poor measures of deontological and consequentialist moral judgment, and so they should not be used to test the mDPM at all (e.g., Rosas and Koenigs, 2014; Kahane, 2015). While I find much of these critiques persuasive, many do not, and sacrificial moral dilemmas remain a common (perhaps the most common) measure of deontological and utilitarian moral judgment in moral psychological research.

from.¹⁰ Fifty-one effect sizes had been observed with sets of just personal dilemmas, while 16 effect sizes had been observed with sets of just impersonal dilemmas. The remaining effect sizes were based on a mix of personal and impersonal sacrificial dilemmas ($k = 42$). The overall effect of dilemma type was not significant: $F_{(3,106)} = 1.47$, $p = 0.227$. Moreover, contrary to expectation, the pooled effect from sets of only personal dilemmas did not differ significantly from zero [$g = 0.03$, 95% CI = (−0.06, 0.13), $p = 0.512$] and was *smaller* than the pooled effect from sets of only impersonal dilemmas or from a mix of the two types (see Figure 5; neither difference was statistically significant).

3.4.6 Person dilemmas: conflict

Koenigs et al. (2007) further split personal dilemmas into low conflict and high conflict dilemmas. In low conflict personal dilemmas, there is complete agreement or almost complete agreement among participants about which choice the agent should take. Conversely, in high conflict dilemmas, there is substantial disagreement among participants about which choice the agent should take. Greene et al. (2008), in addition to recommending that only personal dilemmas be used in tests of the mDPM, further stated that researchers should focus on high conflict dilemmas because “[o]nly high-conflict dilemmas are suitable for examining the conflict between utilitarian [= consequentialist] and non-utilitarian judgment processes” (p. 1148). A large majority (83.3%) of the effect sizes based on just personal dilemmas came from studies that used high conflict dilemmas exclusively.¹¹ The pooled effect of studies with all high conflict personal dilemmas was not significantly different from zero, $g = 0.05$, 95% CI = (−0.04, 0.14), $p = 0.244$, and did not differ significantly from the pooled effect of studies which also featured some low conflict personal dilemmas: $F_{(2,91)} = 0.92$, $p = 0.404$.

3.4.7 Manipulation check

If a study fails to find an effect of a cognitive-processing manipulation on an outcome, one explanation is that cognitive processing does not have an impact on that particular outcome. An alternative explanation, however, is that the cognitive-processing

¹⁰ To do this, I first checked if the authors themselves stated the proportion of personal dilemmas in their study (with reference to Greene et al., 2001, where this the distinction between personal and impersonal dilemmas was first introduced). If that failed, I next checked if one or more of the dilemmas had been part of the dilemma battery used in Greene et al. (2001). If that too failed, I read the dilemma's text in full and then coded it myself based on the three criteria described in Greene et al. (2001).

¹¹ To determine this, I first checked if the authors themselves stated the proportion of high conflict personal dilemmas in their article (with reference to the distinction made in Koenigs et al., 2007). If that failed, I checked if one or more of the dilemmas had been part of the battery of high conflict personal dilemmas used in Koenigs et al. (2007). In this way, I was able to determine the proportion of high conflict personal dilemmas used in all but three reported effect sizes. Since the high conflict/low conflict distinction is an empirical distinction (unlike the personal/impersonal distinction, which is a conceptual distinction), I was not able to categorize these remaining instances myself and so they were excluded from the analysis.

manipulation in question failed to do what it was supposed to do. To help rule out this alternative explanation, manipulation checks are essential (see Horstmann et al., 2010, p. 228–230). Of the effect sizes included in the meta-analysis, 58 came from studies that included at least one manipulation check, while 51 came from studies that did not report any manipulation checks. The most common type of manipulation check compared participant response times across conditions (see Table 1), the thought being that participants who have been encouraged to rely on type 2 processing should take longer to respond than control participants or participants whose ability to engage in type 2 processing has been inhibited. Almost all manipulation checks were reported as successful ($k = 56$). The overall effect of manipulation check status was not significant: $F_{(3,106)} = 1.80$, $p = 0.152$. In particular, the pooled effect of studies with successful manipulation checks did not differ significantly from the other two pooled effects ($F_s \leq 2.68$, $p_s \geq 0.074$)—it did, however, differ significantly from zero, $g = 0.11$, 95% CI = (0.01, 0.21), $p = 0.039$.

3.4.8 Type of literature

Ninety-seven effect sizes came from studies that had undergone peer review; the remaining 12 came from studies in the gray literature. The effect of type of literature was not significant, $F_{(2,107)} = 3.06$, $p = 0.051$. However, it is worth noting that while the pooled effect of peer reviewed studies was significant and in the direction of the mDPM [$g = 0.08$, 95% CI = (0.01, 0.14), $p = 0.023$], the pooled effect of studies from the gray literature was in the opposite direction [though not significant; $g = -0.05$, 95% CI = (-0.24, 0.14), $p = 0.578$]. This further (see Section 3.3) suggests the presence of some amount of publication bias in the peer reviewed literature on the effect of cognitive-processing manipulations on moral judgments about acts or decisions in sacrificial moral dilemmas.

4 Discussion

Greene's influential dual-process model of moral cognition (mDPM; Greene et al., 2001; Greene, 2008, 2014) proposes that when people engage in Type 1 processing, they typically make deontological moral judgments, while when people engage in Type 2 processing, they typically make consequentialist moral judgments. Since the mDPM is a causal model, a convincing case for it requires experimental evidence. Historically, one major source of causal evidence that proponents of the mDPM have cited are studies that investigate the effect of experimental manipulations designed to either encourage or inhibit Type 2 processing on moral judgments about sacrificial moral dilemmas. This paper meta-analyzed published English-language studies with adult participants that used any of four standard types of cognitive-processing manipulations—cognitive load, ego depletion, induction and time restriction [$n = 44$; $k = 68$; total $N = 14,003$; $M(N) = 194.5$]. The results do not support the mDPM. The overall pooled effect, while in the direction of the mDPM's prediction, was very small and did not differ significantly from zero [$g = 0.06$, 95% CI = (0.00, 0.12), $p =$

0.057]. A meta-analysis of a subset of studies [$n = 8$; $k = 12$; total $N = 2,577$; $M(N) = 214.8$] that allowed for analysis using the process dissociation approach (Jacoby et al., 1987; Conway and Gawronski, 2013) also failed to find support for the mDPM.

One possible objection points out that many of effect sizes included in the first meta-analytic estimate came from studies that featured (some only featured) impersonal sacrificial moral dilemmas. However, Greene et al. (2008, p. 1147) argued that only personal sacrificial dilemmas should be used in tests of the mDPM, and so these results do not in fact threaten the mDPM. To address this response, I used meta-regression, but found that the pooled effect of studies that only included personal dilemmas did not differ significantly from zero or the pooled effect of studies that only used impersonal sacrificial dilemmas.

Another response is to argue that my results, instead of challenging the mDPM, speak only against the effectiveness of the cognitive-processing manipulations I included. If these manipulations had often been unsuccessful in encouraging or inhibiting Type 2 processing, then a significant pooled effect would not be expected even if the mDPM were true. However, I found no significant differences between effect sizes from studies with successful manipulation checks, effect sizes from studies with failed manipulation checks and effect sizes from studies where the authors did not report having used a manipulation check. Of course, manipulation checks can be inappropriate or insufficient, and so this result does not entirely rule out the possibility that most or even all studies used ineffective experimental manipulations. In that case, however, these studies would simply be uninformative and so would also not support the mDPM.

It may also be objected that I failed to include certain experimental manipulations and that, if these had been included, the results of the meta-analysis might have been more favorable to the mDPM. One such omission are studies that investigate the so-called foreign language effect. These studies compare moral judgments made in a foreign language with moral judgments made in one's native language (for a meta-analysis, see Circi et al., 2021). However, even though some authors have indeed suggested that this manipulation increases Type 2 processing and so can be used to test the mDPM (e.g., Brouwer, 2019; Circi et al., 2021), the existing evidence does not bear this out. Several studies have concluded that making moral judgments in a foreign language does not encourage Type 2 processing and instead point to increased psychological distance and reduced emotional responding as more likely explanations for the foreign language effect (e.g., Geipel et al., 2015; Corey et al., 2017; Hayakawa et al., 2017).

In their review of cognitive-processing manipulations, Horstmann et al. (2010) discuss another method I did not consider: manipulations of mood. The motivating idea behind these manipulations is that "sad mood leads people to analyze information more deliberately and thoroughly, whereas a happy mood activates more heuristic, intuitive strategies" (p. 227). At least two studies have investigated the effect of mood on moral judgments about sacrificial dilemmas (Valdesolo and DeSteno, 2006; Pastötter et al., 2013). While neither study

unequivocally supports the mDPM¹² and I know of no other studies like them, experimental manipulations of mood were outside the scope of my literature search.¹³ Therefore, for all I know, these studies are out there and their inclusion may have changed the results of the meta-analysis, perhaps to something more favorable to the mDPM. This is a clear limitation of my paper.

That said, defenders of the mDPM should not take too much comfort in this limitation. Recall that none of the four types of cognitive-processing manipulations I did include showed a significant individual meta-analytic effect. Moreover, the pooled effect of studies that used a combination of these types was significantly *smaller* than zero. This is notable, as Horstmann et al. (2010) recommend researchers use a combination of different cognitive-processing manipulations because “additive effects enhance the probability of a successful manipulation” (p. 233). If cognitive load, ego depletion, induction, time restriction and their combinations are typically successful at encouraging or inhibiting Type 2 processing, these results contradict the mDPM. The fact that I omitted experimental manipulations of mood can overturn this conclusion only if there is reason to think that these manipulations are superior to the cognitive-processing manipulations I did include. Otherwise, even if the meta-analytic effect of studies that manipulated mood had been significant and in the direction of the mDPM, the preponderance of the experimental evidence would still fail to support the model.

Yet another response is that the results of this meta-analysis do not threaten the mDPM because sacrificial moral dilemmas are a poor measure of consequentialist moral judgment (cf., Rosas and Koenigs, 2014; Kahane, 2015). Instead, we should look to research with alternative measures, for example, the Oxford Utilitarianism Scale (Kahane et al., 2018; Capraro et al., 2019). While I personally find this response compelling, for defenders of the mDPM, it comes at a very steep price. A large majority of the evidence for the mDPM—not just experimental evidence, but evidence of any kind—comes from studies with sacrificial moral dilemmas (see e.g., Greene, 2014, p. 700–705). Therefore, to admit that these scenarios are a poor measure of consequentialist moral judgment would be to dramatically undercut the model’s empirical support. It is doubtful, then,

12 Valdesolo and DeSteno (2006) found that positive mood increased, rather than decreased, consequentialist responses in Footbridge; Pastötter et al. (2013) found that the effects of positive and negative mood on moral judgments in Footbridge depends on decision frame (active vs. passive).

13 Greene has also argued that deontological (but not consequentialist) moral judgments are based on emotional processes (see text footnote⁴). While logically distinct from the mDPM’s claim that deontological (but not consequentialist) moral judgments are typically the output of Type 1 processing, the two claims are not always clearly disentangled. Studies that investigate the effect of mood on moral judgments about sacrificial dilemmas bear not only on the mDPM, but also on this emotions claim—indeed, this is the context in which Valdesolo and DeSteno (2006) and Pastötter et al. (2013) manipulated mood. As the causal role of the emotions was not the topic of this meta-analysis, I chose to exclude mood manipulations, to avoid an interpretational muddle.

that defenders of the mDPM would want to champion this particular objection.

A final response is to question the usefulness of meta-analyses, more generally. The results of a meta-analysis, so the objection, are only informative if the studies that go into it are of high quality (e.g., Egger et al., 2001; Borenstein et al., 2009). Conversely, if a “meta-analysis includes many low-quality studies, then fundamental errors in the primary studies will be carried over to the meta-analysis” (Borenstein et al., 2009, p. 380)—garbage in, garbage out. When it comes to the studies included in the present meta-analysis, there are indeed a few reasons for concern. Many of these studies featured relatively small samples and did not justify their sample size; almost half (43.5%) did not report having used a manipulation check; and very few studies had been pre-registered. Yet this objection (like the previous objection) again requires defenders of the mDPM to distance themselves from a substantial chunk of the evidence for their model. In other words: If the objection was successful, this victory would very much be a Pyrrhic victory.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://osf.io/cnury/>.

Author contributions

PR: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was part of the ERC-funded project “The Enemy of the Good. Towards a Theory of Moral Progress” (grant number: 851043).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., et al. (2023). *markdown: Dynamic Documents for R*. Available online at: <https://github.com/rstudio/markdown>
- Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *Am. Psychol.* 63, 602–614. doi: 10.1037/0003-066X.63.7.602
- Awad, E., Dsouza, S., Shariff, A., Rahwan, I., and Bonnefon, J.-F. (2020). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proc. Nat. Acad. Sci. U. S. A.* 117, 2332–2337. doi: 10.1073/pnas.1911517117
- *Bago, B., and De Neys, W. (2019). The intuitive greater good: testing the corrective dual process model of moral cognition. *J. Exp. Psychol. Gen.* 148, 1782–1801. doi: 10.1037/xge0000533
- *Bialek, M., and De Neys, W. (2017). Dual processes and moral conflict: evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgm. Decis. Mak.* 12, 148–167. doi: 10.1017/S1930297500005696
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester: Wiley.
- Brouwer, S. (2019). The auditory foreign-language effect of moral decision making in highly proficient bilinguals. *J. Multiling. Multicult. Dev.* 40, 865–878. doi: 10.1080/01434632.2019.1585863
- Capraro, V., Everett, J. A., and Earp, B. D. (2019). Priming intuition disfavors instrumental harm but not impartial beneficence. *J. Exp. Soc. Psychol.* 83, 142–149. doi: 10.1016/j.jesp.2019.04.006
- Carter, E. C., Schönbrodt, F. D., Gervais, W. M., and Hilgard, J. (2019). Correcting for bias in psychology: a comparison of meta-analytic methods. *Adv. Methods Pract. Psychol. Sci.* 2, 115–144. doi: 10.1177/2515245919847196
- *Cellini, N., Lotto, L., Pletti, C., and Sarlo, M. (2017). Daytime REM sleep affects emotional experience but not decision choices in moral dilemmas. *Sci. Rep.* 7:11059. doi: 10.1038/s41598-017-11530-4
- *Chaturapanich, T., and Chaiwitkornwanich, A. (2015). Belief in a just world and judgment in moral dilemmas. *Asian Soc. Sci.* 11, 178–191. doi: 10.5539/ass.v11n23p178
- Circi, R., Gatti, D., Russo, V., and Vecchi, T. (2021). The foreign language effect on decision-making: a meta-analysis. *Psychon. Bull. Rev.* 28, 1131–1141. doi: 10.3758/s13423-020-01871-z
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd Edn*. Hillsdale, NJ: L. Erlbaum Associates.
- Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., and Garcia-Retamero, R. (2012). Measuring risk literacy: the Berlin numeracy test. *Judgm. Decis. Mak.* 7, 25–47. doi: 10.1017/S1930297500001819
- *Conway, P., and Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. *J. Pers. Soc. Psychol.* 104, 216–235. doi: 10.1037/a0031021
- Corey, J. D., Hayakawa, S., Foucart, A., Aparici, M., Botella, J., Costa, A., et al. (2017). Our moral choices are foreign to us. *J. Exp. Psychol. Learn. Mem. Cognit.* 43, 1109–1128. doi: 10.1037/xlm0000356
- *Cova, F., Strickland, B., Abatista, A., Allard, A., Andow, J., Attie, M., et al. (2018). Estimating the reproducibility of experimental philosophy. *Rev. Philos. Psychol.* 12, 9–44. doi: 10.1007/s13164-018-0407-2
- *Cummins, D. D., and Cummins, R. C. (2012). Emotion and deliberative reasoning in moral judgment. *Front. Psychol.* 3:328. doi: 10.3389/fpsyg.2012.00328
- *Doerflinger, J. T., and Gollwitzer, P. M. (2020). Emotion emphasis effects in moral judgment are moderated by mindsets. *Motiv. Emot.* 44, 880–896. doi: 10.1007/s11031-020-09847-1
- Duval, S., and Tweedie, R. (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56, 455–463. doi: 10.1111/j.0006-341X.2000.00455.x
- Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315, 629–634. doi: 10.1136/bmj.315.7109.629
- Egger, M., Smith, G. D., and Sterne, J. A. C. (2001). Uses and abuses of meta-analysis. *Clin. Med.* 1, 478–484. doi: 10.7861/clinmedicine.1-6-478
- Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.* 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629
- Evans, J., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspect. Psychol. Sci.* 8, 223–241. doi: 10.1177/1745691612460685
- Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics* 90, 891–904. doi: 10.1007/s11192-011-0494-7
- Ferguson, C. J., and Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychol. Methods* 17, 120–128. doi: 10.1037/a0024445
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Ox. Rev.* 5, 5–15.
- Franco, A., Malhotra, N., and Simonovits, G. (2014). Publication bias in the social sciences: unlocking the file drawer. *Science* 345, 1502–1505. doi: 10.1126/science.1255484
- *Gamez-Djokic, M., and Molden, D. (2016). Beyond affective influences on deontological moral judgment: the role of motivations for prevention in the moral condemnation of harm. *Pers. Soc. Psychol. Bull.* 42, 1522–1537. doi: 10.1177/0146167216665094
- *Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., and Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: the CNI model of moral decision-making. *J. Pers. Soc. Psychol.* 113, 343–376. doi: 10.1037/pspa0000086
- Geipel, J., Hadjichristidis, C., and Surian, L. (2015). The foreign language effect on moral judgment: the role of emotions and norms. *PLoS ONE* 10:e0131529. doi: 10.1371/journal.pone.0131529
- Gilbert, D. T., Pelham, B. W., and Krull, D. S. (1988). On cognitive busyness: when person perceivers meet persons perceived. *J. Pers. Soc. Psychol.* 54, 733–740. doi: 10.1037/0022-3514.54.5.733
- Greene, J. D. (2008). “The secret joke of Kant’s soul,” in *Moral Psychology: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. W. Sinnott-Armstrong (Cambridge, MA: MIT Press), 35–79.
- Greene, J. D. (2014). Beyond point-and-shoot morality: why cognitive (neuro)science matters for ethics. *Ethics* 124, 695–726. doi: 10.1086/675875
- *Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., and Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition* 107, 1144–1154. doi: 10.1016/j.cognition.2007.11.004
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., and Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400. doi: 10.1016/j.neuron.2004.09.027
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108. doi: 10.1126/science.1062872
- Guglielmo, S. (2015). Moral judgment as information processing: an integrative review. *Front. Psychol.* 6:1637. doi: 10.3389/fpsyg.2015.01637
- *Gürçay, B., and Baron, J. (2017). Challenges for the sequential two-system model of moral judgement. *Think. Reason.* 23, 49–80. doi: 10.1080/13546783.2016.1216011
- Hagger, M. S., Wood, C., Stiff, C., and Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychol. Bull.* 136, 495–525. doi: 10.1037/a0019486
- *Ham, J., and van den Bos, K. (2010). On unconscious morality: the effects of unconscious thinking on moral decision making. *Soc. Cogn.* 28, 74–83. doi: 10.1521/soco.2010.28.1.74
- Harrer, M., Cuijpers, P., Furukawa, T. A., and Ebert, D. D. (2021). *Doing Meta-Analysis with R: A Hands-on Guide*. Boca Raton, FL: CRC Press.
- Harzing, A. W. (2021). *Publish or Perish*. Available online at: <https://harzing.com/resources/publish-or-perish>
- *Hashimoto, H., Maeda, K., and Matsumura, K. (2022). Fickle judgments in moral dilemmas: time pressure and utilitarian judgments in an interdependent culture. *Front. Psychol.* 13:795732. doi: 10.3389/fpsyg.2022.795732
- Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., and Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind Lang.* 22, 1–21. doi: 10.1111/j.1468-0017.2006.00297.x
- Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., and Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychol. Sci.* 28, 1387–1397. doi: 10.1177/0956797617720944
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83. doi: 10.1017/S0140525X0999152X
- Horstmann, N., Hausmann, D., and Ryf, S. (2010). “Methods for inducing intuitive and deliberate processing modes,” in *Foundations for Tracing Intuition: Challenges and Methods*, eds A. Glöckner, and C. Wittmann (New York, NY: Psychology Press), 219–237.
- *Houston, C. J. (2010). *Short-Term Retention, Time Pressure, and Accessibility Tasks Do Not Interfere with Utilitarian Moral Judgment* (Master’s thesis). Harvard University.
- Isler, O., and Yilmaz, O. (2023). How to activate intuitive and reflective thinking in behavior research? A comprehensive examination of experimental techniques. *Behav. Res. Methods* 55, 3679–3698. doi: 10.3758/s13428-022-01984-4
- Jacoby, J., Jaccard, J., Kuss, A., Troutman, T., and Mazursky, D. (1987). New directions in behavioral process research: implications for social psychology. *J. Exp. Soc. Psychol.* 23, 146–175. doi: 10.1016/0022-1031(87)90029-1

- Kahane, G. (2012). On the wrong track: process and content in moral psychology: process and content in moral psychology. *Mind Lang.* 27, 519–545. doi: 10.1111/mila.12001
- Kahane, G. (2015). Sidetracked by trolleys: why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. *Soc. Neurosci.* 10, 551–560. doi: 10.1080/17470919.2015.1023400
- Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., et al. (2018). Beyond sacrificial harm: a two-dimensional model of utilitarian psychology. *Psychol. Rev.* 125, 131–164. doi: 10.1037/rev0000093
- *Killgore, W. D., Killgore, D. B., Day, L. M., Li, C., Kamimori, G. H., and Balkin, T. J. (2007). The effects of 53 hours of sleep deprivation on moral judgment. *Sleep* 30, 345–352. doi: 10.1093/sleep/30.3.345
- Kinnunen, S. P., and Windmann, S. (2013). Dual-processing altruism. *Front. Psychol.* 4:193. doi: 10.3389/fpsyg.2013.00193
- Klenk, M. (2022). The influence of situational factors in sacrificial dilemmas on utilitarian moral judgments: a systematic review and meta-analysis. *Rev. Philos. Psychol.* 13, 593–625. doi: 10.1007/s13164-021-00547-4
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446, 908–911. doi: 10.1038/nature05631
- *Körner, A., and Volk, S. (2014). Concrete and abstract ways to deontology: cognitive capacity moderates construal level effects on moral judgments. *J. Exp. Soc. Psychol.* 55:139–145. doi: 10.1016/j.jesp.2014.07.002
- *Kroneisen, M., and Steghaus, S. (2021). The influence of decision time on sensitivity for consequences, moral norms, and preferences for inaction: time, moral judgments, and the CNI model. *J. Behav. Decis. Mak.* 34, 140–153. doi: 10.1002/bdm.2202
- Kühberger, A., Fritz, A., and Scherndl, T. (2014). Publication bias in psychology: a diagnosis based on the correlation between effect size and sample size. *PLoS ONE* 9:e105825. doi: 10.1371/journal.pone.0105825
- Kvarven, A., Strömmland, E., Wollbrant, C., Andersson, D., Magnus, J. Johannesson, Tinghög, G., Västfjäll, D., et al. (2020). The intuitive cooperation hypothesis revisited: a meta-analytic examination of effect size and between-study heterogeneity. *J. Econ. Sci. Assoc.* 6, 26–42. doi: 10.1007/s40881-020-00084-3
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863
- Lane, D., and Sulikowski, D. (2017). Bleeding-heart conservatives and hard-headed liberals: the dual processes of moral judgements. *Pers. Individ. Dif.* 115, 30–34. doi: 10.1016/j.paid.2016.09.045
- *Li, Z., Gao, L., Zhao, X., and Li, B. (2021). Deconfounding the effects of acute stress on abstract moral dilemma judgment. *Curr. Psychol.* 40, 5005–5018. doi: 10.1007/s12144-019-00453-0
- *Li, Z., Xia, S., Wu, X., and Chen, Z. (2018). Analytical thinking style leads to more utilitarian moral judgments: an exploration with a process-dissociation approach. *Pers. Individ. Dif.* 131, 180–184. doi: 10.1016/j.paid.2018.04.046
- *Liu, C., and Liao, J. (2021). Stand up to action: the postural effect of moral dilemma decision-making and the moderating role of dual processes. *PsyCh J.* 10, 587–597. doi: 10.1002/pchj.449
- Lüdtke, D. (2019). *esc: Effect Size Computation for Meta Analysis*. Available online at: <https://CRAN.R-project.org/package=esc>
- *Lyrantzis, E. A. (2017). *Calibrating the Moral Compass: The Effect of Tailored Communications on Non-Profit Advertising* (Doctoral thesis). Claremont Graduate University.
- McDonald, K., Graves, R., Yin, S., Weese, T., and Sinnott-Armstrong, W. (2021). Valence framing effects on moral judgments: a meta-analysis. *Cognition* 212:104703. doi: 10.1016/j.cognition.2021.104703
- *McPhetres, J., Conway, P., Hughes, J. S., and Zuckerman, M. (2018). Reflecting on god's will: reflective processing contributes to religious peoples' deontological dilemma responses. *J. Exp. Soc. Psychol.* 79, 301–314. doi: 10.1016/j.jesp.2018.08.013
- Medin, D., Ojalehto, B., Marin, A., and Bang, M. (2017). Systems of (non-)diversity. *Nat. Hum. Behav.* 1:0088. doi: 10.1038/s41562-017-0088
- Morris, S. B., and DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol. Methods* 7, 105–125. doi: 10.1037/1082-989X.7.1.105
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021a). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 372:n160. doi: 10.1136/bmj.n160
- Page, M. J., Sterne, J. A. C., Higgins, J. P. T., and Egger, M. (2021b). Investigating and dealing with publication bias and other reporting biases in meta-analyses of health research: a review. *Res. Synth. Methods* 12, 248–259. doi: 10.1002/jrsm.1468
- Pastötter, B., Gleixner, S., Neuhauser, T., and Bäuml, K.-H. T. (2013). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition* 126, 373–377. doi: 10.1016/j.cognition.2012.11.003
- Patil, I., Zucchelli, M. M., Kool, W., Campbell, S., Fornasier, F., Calò, M., et al. (2021). Reasoning supports utilitarian resolutions to moral dilemmas across diverse measures. *J. Pers. Soc. Psychol.* 120, 443–460. doi: 10.1037/pspp0000281
- *Paxton, J. M., Bruni, T., and Greene, J. D. (2014). Are 'counter-intuitive' deontological judgments really counter-intuitive? An empirical reply to Kahane et al. (2012). *Soc. Cogn. Affect. Neurosci.* 9, 1368–1371. doi: 10.1093/scan/nst102
- *Paxton, J. M., Ungar, L., and Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cogn. Sci.* 36, 163–177. doi: 10.1111/j.1551-6709.2011.01210.x
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., and Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J. Clin. Epidemiol.* 61, 991–996. doi: 10.1016/j.jclinepi.2007.11.010
- Peterson, R. A. (2001). On the use of college students in social science research: insights from a second-order meta-analysis. *J. Cons. Res.* 28, 450–461. doi: 10.1086/323732
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/>
- Rad, M. S., Martingano, A. J., and Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proc. Nat. Acad. Sci. U. S. A.* 115, 11401–11405. doi: 10.1073/pnas.1721165115
- Rand, D. G. (2016). Cooperation, fast and slow: meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychol. Sci.* 27, 1192–1206. doi: 10.1177/0956797616654455
- *Reid, E. (2018). *Moral Decision Making* (Bachelor thesis). University of Tasmania.
- *Rosas, A., and Aguilar-Pardo, D. (2020). Extreme time-pressure reveals utilitarian intuitions in sacrificial dilemmas. *Think. Reason.* 26, 534–551. doi: 10.1080/13546783.2019.1679665
- *Rosas, A., Bermúdez, J. P., and Aguilar-Pardo, D. (2019). Decision conflict drives reaction times and utilitarian responses in sacrificial dilemmas. *Judgm. Decis. Mak.* 14, 555–564. doi: 10.1017/S193029750000485X
- Rosas, A., and Koenigs, M. (2014). Beyond "utilitarianism": maximizing the clinical impact of moral judgment research. *Soc. Neurosci.* 9, 661–667. doi: 10.1080/17470919.2014.937506
- Rothstein, H., Sutton, A. J., and Borenstein, M. (2005). "Publication bias in meta-analysis," in *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, eds H. Rothstein, A. J. Sutton, and M. Borenstein (Chichester; Hoboken, NJ: Wiley), 1–7.
- Schein, C., and Gray, K. (2018). The theory of dyadic morality: reinventing moral judgment by redefining harm. *Person. Soc. Psychol. Rev.* 22, 32–70. doi: 10.1177/1088868317698288
- *Schwitzgebel, E., and Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection. *Cognition* 141, 127–137. doi: 10.1016/j.cognition.2015.04.015
- Simpson, D. (2021). *The Dual-Process Model and Moral Dilemmas: Reflection Does Not Drive Self-Sacrifice* (Master's thesis). Georgia State University.
- *Spears, D., Okan, Y., Hinojosa-Aguayo, I., Perales, J. C., Ruz, M., and González, F. (2021). Can induced reflection affect moral decision-making? *Philos. Psychol.* 34, 28–46. doi: 10.1080/09515089.2020.1861234
- Stanley, T. D., and Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Res. Synth. Methods* 5, 60–78. doi: 10.1002/jrsm.1095
- *Starcke, K., Ludwig, A.-C., and Brand, M. (2012). Anticipatory stress interferes with utilitarian moral judgment. *Judgm. Decis. Mak.* 7, 61–68. doi: 10.1017/S1930297500001832
- Stewart, N., Chandler, J., and Paolacci, G. (2017). Crowdsourcing samples in cognitive science. *Trends Cogn. Sci.* 21, 736–748. doi: 10.1016/j.tics.2017.06.007
- *Suter, R. S., and Hertwig, R. (2011). Time and moral judgment. *Cognition* 119, 454–458. doi: 10.1016/j.cognition.2011.01.018
- *Swann, W. B., Gómez, Á., Buhrmester, M. D., López-Rodríguez, L., Jiménez, J., and Vázquez, A. (2014). Contemplating the ultimate sacrifice: identity fusion channels group affect, cognition, and moral decision making. *J. Pers. Soc. Psychol.* 106, 713–727. doi: 10.1037/a0035809
- Thalmayer, A. G., Toscanelli, C., and Arnett, J. J. (2021). The neglected 95% revisited: is American psychology becoming less American? *Am. Psychol.* 76, 116–129. doi: 10.1037/amp0000622
- Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *Monist* 59, 204–217. doi: 10.5840/monist197659224
- Thornton, A., and Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *J. Clin. Epidemiol.* 53, 207–216. doi: 10.1016/S0895-4356(99)00161-4

- *Timmons, S., and Byrne, R. M. (2019). Moral fatigue: The effects of cognitive fatigue on moral reasoning. *Q. J. Exp. Psychol.* 72, 943–954. doi: 10.1177/1747021818772045
- *Tinghög, G., Andersson, D., Bonn, C., Johannesson, M., Kirchler, M., Koppel, L., et al. (2016). Intuition and moral decision-making—the effect of time pressure and cognitive load on moral judgment and altruistic behavior. *PLoS ONE* 11:e0164012. doi: 10.1371/journal.pone.0164012
- *Trémolière, B., and Bonnefon, J.-F. (2014). Efficient kill-save ratios ease up the cognitive demands on counterintuitive moral utilitarianism. *Pers. Soc. Psychol. Bull.* 40, 923–930. doi: 10.1177/0146167214530436
- *Trémolière, B., Neys, W. D., and Bonnefon, J.-F. (2012). Mortality salience and morality: thinking about death makes people less utilitarian. *Cognition* 124, 379–384. doi: 10.1016/j.cognition.2012.05.011
- Valdesolo, P., and DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychol. Sci.* 17, 476–477. doi: 10.1111/j.1467-9280.2006.01731.x
- *Van 't Veer, A. E., and Slegers, W. W. (2019). Psychology data from an exploration of the effect of anticipatory stress on disgust vs. non-disgust related moral judgments. *J. Open Psychol. Data* 7, 1–6. doi: 10.5334/jopd.43
- *Vega, S., Mata, A., Ferreira, M. B., and Vaz, A. R. (2021). Metacognition in moral decisions: judgment extremity and feeling of rightness in moral intuitions. *Think. Reason.* 27, 124–141. doi: 10.1080/13546783.2020.1741448
- Vicario, C. M., Kuran, K. A., Rogers, R., and Rafal, R. D. (2018). The effect of hunger and satiety in the judgment of ethical violations. *Brain Cogn.* 125, 32–36. doi: 10.1016/j.bandc.2018.05.003
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor Package. *J. Stat. Soft.* 36, 1–48. doi: 10.18637/jss.v036.i03
- *Weippert, M., Rickler, M., Kluck, S., Behrens, K., Bastian, M., Mau-Moeller, A., et al. (2018). It's harder to push, when I have to push hard—physical exertion and fatigue changes reasoning and decision-making on hypothetical moral dilemmas in males. *Front. Behav. Neurosci.* 12:268. doi: 10.3389/fnbeh.2018.00268
- Wickham, H. (2022). *stringr: Simple, Consistent Wrappers for common String Operations*. Available online at: <https://CRAN.R-project.org/package=Stringr>
- Wickham, H., François, R., Henry, L., Müller, K., and Vaughan, D. (2023a). *dplyr: A Grammar of Data Manipulation*. Available online at: <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., Hester, J., and Bryan, J. (2023b). *readr: Read Rectangular Text Data*. Available online at: <https://CRAN.R-project.org/package=readr>
- Wright, P. (1974). The harassed decision maker: time pressures, distractions, and the use of evidence. *J. Appl. Psychol.* 59, 555–561. doi: 10.1037/h0037186
- Xie, Y. (2023). *bookdown: Authoring Books and Technical Documents with R Markdown*. Available online at: <https://github.com/rstudio/bookdown>
- *Youssef, F. F., Dookeeram, K., Basdeo, V., Francis, E., Doman, M., Mamed, D., et al. (2012). Stress alters personal moral decision making. *Psychoneuroendocrinology* 37, 491–498. doi: 10.1016/j.psyneuen.2011.07.017
- *Zhao, J., Harris, M., and Vigo, R. (2016). Anxiety and moral judgment: the shared deontological tendency of the behavioral inhibition system and the unique utilitarian tendency of trait anxiety. *Pers. Individ. Dif.* 95, 29–33. doi: 10.1016/j.paid.2016.02.024
- *Zheng, M. (2018). *How Power Affects Moral Judgment* (Doctoral thesis). University College London.

*Studies included in the meta-analysis.