Judgment and Decision Making (2026), **21**:e30 1–24
doi:10.1017/jdm.2026.10051

EMPIRICAL ARTICLE

Cognitive reflection predicts utilitarian behavior in the monetary moral dilemma after 17 months

Rozelin Vurgun[1], Ensar Acem[1], Fırat Seker[1,2], Burak Dogruyol[1], Hasan G. Bahcekapili[3], Ozan Isler[4], and Onurcan Yilmaz[1]

[1]Department of Psychology, Kadir Has University, Istanbul, Türkiye; [2]Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, USA; [3]Department of Psychology, Istanbul Medipol University, Istanbul, Türkiye and [4]School of Economics, The University of Queensland, Brisbane, Australia

Corresponding author: Rozelin Vurgun (rozelin.vurgun@khas.edu.tr) and Onurcan Yilmaz (onurcan.yilmaz@khas.edu.tr)

Received: 29 October 2025 UTC; **Revised:** 2 September 2026 UTC; **Accepted:** 3 September 2026 UTC

Keywords: utilitarianism; cognitive reflection; dual-process model; impartial beneficence; instrumental harm

Abstract

The dual-process model of moral judgment proposes that utilitarian choices arise from reflective processes that can override intuitive, rule-based responses. Yet most supporting evidence relies on sacrificial dilemmas with limited ecological validity and problematic associations with antisocial traits. This preregistered study investigated whether individual differences in reflection predict utilitarian behavior in an incentivized setting. At Time 1, 1,200 U.S. participants completed performance-based measures of reflection, reasoning ability, dispositional thinking styles, and the Oxford Utilitarianism Scale (OUS). Seventeen months later, 321 participants completed a monetary moral dilemma with real economic consequences and reported their moral justifications. Initial measures of reflection, including the Cognitive Performance Test, Raven's matrices, and open-minded thinking, predicted utilitarian behavior. Neither impartial beneficence nor instrumental harm subscales of the OUS predicted utilitarian behavior. Exploratory analyses showed that religiosity moderated the reflection–utilitarian behavior association: reflection predicted utilitarian behavior more strongly among non-believers than believers. These findings highlight reflection as a predictor of utilitarian behavior across time, question the validity of self-report utilitarianism scales such as the OUS, and underscore the importance of religious belief in cognitive reflection when examining individual differences in moral behavior.

1. Introduction

Understanding how people arrive at moral judgments has long been a central concern of psychology. Dual-process accounts of moral cognition argue that intuitive processes are closely tied to deontological judgments grounded in rules and duties, while reflective processes promote utilitarian judgments that prioritize outcomes (Cushman & Greene, 2012; Greene, 2008; Greene et al., 2001, 2004). Much of the evidence supporting this model comes from sacrificial trolley-type dilemmas in which harming one person to save several others is treated as utilitarian reasoning (Greene & Haidt, 2002; Patil et al., 2021). However, such scenarios have been criticized for their limited external validity and conceptual ambiguity (Bartels & Pizarro, 2011; Bauman et al., 2014; Crone & Laham, 2017; Kahane et al., 2015; Körner & Deutsch, 2022). Moreover, meta-analyses that pool small sample-size experimental studies reveal either no effect or only small effects of reflection manipulations on utilitarian responding,

© The Author(s), 2026. Published by Cambridge University Press on behalf of Society for Judgment and Decision Making and European Association for Decision Making. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

raising questions about the robustness of these findings (Capraro, 2024; Fahrenwaldt et al., 2025; Rehren, 2024). These debates raise a fundamental question: does reflection, when measured as a stable individual difference, predict actual utilitarian behavior in real decision-making contexts? The present research provides one of the first systematic tests of this question, integrating cognitive, dispositional, and self-report measures of reflection with an incentivized behavioral task conducted 17 months later.

The dual-process model remains highly influential in explaining moral judgment. According to this framework, reflective thought is required to override the immediate emotional pull of deontological rules and to prioritize aggregate welfare (Greene, 2008). Evidence in support of this claim has often relied on sacrificial dilemmas, but these paradigms raise two challenges. First, they are highly artificial and may not capture real-world trade-offs (Bauman et al., 2014; Bostyn et al., 2018; Körner & Deutsch, 2022). Second, utilitarian responses in these dilemmas correlate positively with psychopathic traits (Djeriouat & Trémolière, 2014; Kahane et al., 2015; Paruzel-Czachura & Farny, 2023), suggesting they may reflect reduced sensitivity to harm rather than genuine concern for maximizing outcomes (Kahane et al., 2015).

To address these concerns regarding sacrificial dilemmas, researchers have developed more realistic alternatives, including historically-based scenarios (Körner & Deutsch, 2022) and lab-based experiments on real-world behavior (Bostyn et al., 2018; Gold et al., 2015). A particularly promising direction involves using economic games to measure moral behavior directly, thereby avoiding the intention-behavior gap common in self-report studies (Clarkson, 2022). For instance, Bostyn et al. (2019) introduced a monetary trolley dilemma in which participants could reduce one person's payoff to increase those of five others. Unlike sacrificial dilemmas, utilitarian choices in this paradigm were not associated with psychopathy, indicating that monetary trade-offs may better capture impartial, outcome-focused reasoning. Still, much of this work has relied on hypothetical scenarios (e.g., Côté et al., 2013) or designs involving deception (e.g., Bostyn et al., 2019).

Alongside these behavioral innovations, theoretical work has refined our understanding of what constitutes utilitarian decision-making. Kahane et al. (2018) distinguish between instrumental harm, endorsing harm for the greater good, and impartial beneficence, the impartial promotion of overall welfare. This distinction helps explain why sacrificial dilemmas primarily capture instrumental harm. The Oxford Utilitarianism Scale (OUS) has since become a widely used measure of these two dimensions of utilitarianism at the level of individual differences, supported by evidence of satisfactory psychometric properties (Ni et al., 2023), cross-cultural reliability (Oshiro et al., 2024), and even heritability (Bates, 2025). Yet the predictive validity of these self-reported dimensions for real-world utilitarian behavior remains uncertain, a gap this study seeks to address.

Although trolley dilemmas are commonly used to capture utilitarian and deontological moral principles, different moral considerations play roles in moral judgments (Aktas et al., 2017). For instance, virtue-ethical concerns also guide moral decisions when individuals act in accordance with character ideals rather than with outcome- or rule-based reasoning (Uhlmann et al., 2015). While the deontological response in moral dilemmas is typically characterized as not harming others, fatalism may also play a role, whereby individuals avoid moral responsibility by attributing outcomes to fate or chance (Bostyn et al., 2025). In addition, responding to moral dilemmas may not even be morally grounded at all, reflecting amorality (Aktas et al., 2017). Finally, fairness concerns are also considered fundamental in moral decision-making (Fehr & Schmidt, 1999). Especially in monetary contexts, redistributing resources might motivate a focus on equality rather than maximizing overall welfare. Therefore, capturing moral justifications provides a more complete picture of moral reasoning and also serves as a validity check on which moral considerations the dilemma actually elicits.

A second issue concerns measurement of reflection. Experimental manipulations often yield small effects (Capraro, 2024; Fahrenwaldt et al., 2025), but individual-difference measures may provide more reliable insights. Performance-based tests such as the Cognitive Reflection Test (CRT; Frederick, 2005; Sirota & Juanchich, 2018), belief-bias and conjunction-fallacy items (Kahneman & Tversky, 1973; Markovits & Nantel, 1989), reasoning ability (Raven, 2000), and thinking-style questionnaires (Newton et al., 2024) have all been used to capture reflective tendencies. Recent integrative efforts

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

combine these measures into composite indices of reflective cognition (Isler & Yilmaz, 2023). Such approaches allow researchers to test whether reflection, assessed across multiple domains, predicts moral behavior over time.

Finally, reflection must be understood within its sociocultural context, as it is not inherently accuracy-promoting but shaped by epistemic norms (Baron, 2020; Saribay & Yilmaz, 2024; Yilmaz, 2021). Religiosity is especially relevant in this regard. Religious individuals are more likely to endorse deontological reasoning and less likely to make utilitarian judgments in sacrificial moral dilemmas (Antonenko Young et al., 2013; Saroglou & Craninx, 2021). When utilitarian and deontological judgments are examined separately, reflection increases deontological judgments only among more religious individuals, who appear to rely on cognitive resources to access and apply religious moral precepts, while leaving unchanged the deontological responses of less religious individuals (McPhetres et al., 2018). Moreover, religiosity is positively associated with impartial beneficence but not instrumental harm, challenging the previously reported negative link between religiosity and utilitarian judgments involving sacrifice (Kahane et al., 2018; Paruzel-Czachura & Charzyńska, 2023). Epistemically, believers often treat divine authority as a source of moral truth (Baron et al., 2015; Piazza & Landy, 2013; Simpson et al., 2016), whereas non-believers tend to adopt evidence-based reasoning and belief updating (Metz et al., 2018; Yilmaz & Isler, 2019). Reflection, moreover, is not inherently critical: it can be recruited to bolster an already favored conclusion rather than to evaluate it, and believers who regard divine authority as settling moral questions may place less value on such critical evaluation (Baron, 2020; Yilmaz & Isler, 2025). Thus, reflection may promote utilitarian reasoning among non-believers but reinforce rule-based convictions among believers. Despite theoretical arguments and findings based on hypothetical sacrificial dilemmas, no prior research has empirically tested whether religiosity moderates the reflection–utilitarianism link in incentivized behavioral contexts.

1.1. The present research

The present study examined whether reflection predicts utilitarian moral behavior in an externally valid, incentivized setting. Reflection was assessed through performance-based tasks (Isler & Yilmaz, 2023), reasoning ability (Raven, 2000), and dispositional thinking styles (Newton et al., 2024). Seventeen months later, participants completed a monetary moral dilemma (Bostyn et al., 2019) with real economic consequences. We also tested whether utilitarian dimensions measured by the OUS (Kahane et al., 2018) predicted behavior and whether religiosity moderated the reflection–utilitarianism link.

We preregistered three hypotheses: (H1) impartial beneficence and (H2) instrumental harm would correlate positively with utilitarian behavior, and (H3) higher cognitive performance scores would predict a greater likelihood of utilitarian choices, consistent with the dual-process model (Greene, 2008). Moral justifications were preregistered as exploratory variables, and their correlations with utilitarian behavior and cognitive reflection were examined. In addition, we conducted exploratory analyses using theoretically relevant variables available in the Time 1 dataset, including additional cognitive measures (i.e., Raven's Matrices and self-reported thinking styles assessed with the 4-Component Thinking Style Questionnaire) and a single-item measure of religiosity.

This design contributes to the literature by involving real incentives, reducing demand effects and enhancing external validity. It further evaluates the predictive validity of the OUS on utilitarian behavior, integrates multiple measures of reflection to provide a robust test of dual-process theory, and investigates religiosity as a potential moderator of the reflection–utilitarianism association.

2. Method

The research protocol was preregistered on the Open Science Framework before data collection at Time 2 (https://osf.io/skjmh/?view_only=e6f6972639d642999762eb1e3a1866b0).

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

2.1. Participants

A total of 1,200 U.S. participants were recruited at Time 1 (575 females; M_{age} = 39.9, SD_{age} = 14.9) and received $2.00 for survey completion. Seventeen months later, participants were re-contacted for Time 2, yielding 322 responses. As preregistered, one duplicate was excluded, leaving a final sample of 321 participants (161 females; M_{age} = 47.12, SD_{age} = 14.17), with a $2 base payment.

The data were drawn from a larger project on moral beliefs among religious believers and non-believers[1] (the main study's OSF page: https://osf.io/rnzy9/?view_only=e3ca7f26e2d443f396f46892a1e9781b). Participants were recruited via Prolific with preselection for gender balance and equal numbers of Christian believers and non-believers (agnostics/atheists) at Time 1 (N = 1,200). Those identifying as 'other' or non-Christian were excluded in line with the project focus, and 1,169 eligible participants were invited to Time 2 after 17 months.[2] At Time 2, we aimed to balance subgroups by belief (believer and nonbeliever), gender (male and female), and socioeconomic status (high and low); crossing these three binary factors (e.g., male high-socioeconomic status (SES) believers, female low-SES non-believers) yielded eight subgroups in total, with 40 participants targeted per subgroup recruited via Prolific's demographic prescreening criteria (N = 320). We also collected self-reported demographics, including religious affiliation, religiosity, subjective SES, and political ideology. Similar to the targeted prescreening via Prolific, religious affiliation was 49.5% Christian believers, 49.3% non-believers (27.1% atheists, 21.1% agnostics), and 1.2% Jewish/other. Religiosity (0 = not at all religious, 10 = very religious) averaged 3.31 (SD = 3.70). Subjective SES was measured with a 1–10 socioeconomic ladder item, where higher scores indicate higher perceived status in U.S. society (M = 5.30, SD = 1.83). This measure aligned with Prolific's prescreening SES criteria, which also used a 1–10 scale with scores below 5 indicating low SES and scores of 5 and above indicating high SES; accordingly, 49.1% of the sample were classified as low SES and 50.9% as high SES.

2.2. Procedure and materials

Participants' scores on the OUS, Cognitive Performance Task (CPT), 4-Component Thinking Styles Questionnaire (4-CTSQ), Raven's matrices, and religiosity were drawn from the Time 1 data collection. At Time 2, conducted 17 months later, they completed the monetary moral dilemma and rated their decision justifications based on various normative moral frameworks as a validation of the task. The study materials, analysis code, data, and codebook are available on the OSF project site: https://osf.io/6r7gn/?view_only=828e9bf46ac74907b7878e106a8f5b59.

Since this study's data were from a larger project, participants also completed additional scales at Time 1 and economic games at Time 2, with all measures presented in random order. A demographic questionnaire was given at the end of the study (see the Appendix). At the start of Time 2, participants were informed that one randomly selected task would determine their bonus payment, ensuring that their decisions were real and consequential.

2.2.1. The Oxford Utilitarianism Scale (OUS)

The OUS (Kahane et al., 2018) measures the two dimensions of utilitarian thinking. Impartial beneficence represents the positive dimension of utilitarianism, reflecting the principle of maximizing overall well-being with equal concern for all, regardless of parochial ties. The five-item subscale (α = .73) includes items such as, 'It is just as wrong to fail to help someone as it is to actively harm them yourself'. The instrumental harm dimension, on the other hand, represents the negative component of

[1]Time 1 also contained a broad psychological and attitudinal battery. Time 2 included behavioral tasks in addition to the monetary moral dilemma. These materials and data have been published on the main study's OSF page.

[2]The timing of these intervals was determined by the practical constraints of the larger longitudinal project, including a three-wave design, intended to minimize carryover effects from the repeated behavioral moral tasks. The second wave was collected 10 months after the first, and the third 7 months after the second. The monetary moral dilemma was introduced exclusively at the third wave. For the current study examining cognitive reflection and utilitarian behavior, only the first and third waves were used, referred to here as Time 1 and Time 2 for simplicity.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

utilitarianism, reflecting the acceptance of harm to achieve the greater good. The four-item subscale (α = .76) includes items such as, 'It is morally right to harm an innocent person if harming them is a necessary means to helping several other innocent people'. Participants rated their agreement with the items on a Likert scale from 1 (strongly disagree) to 7 (strongly agree).

2.2.2. Cognitive performance task (CPT)

Isler and Yilmaz (2023) introduced the CPT by combining existing reflective and intuitive thinking measures. The first three of its five multiple-choice items are from the validated four-option version of the Cognitive Reflection Test (CRT; Frederick, 2005; Sirota & Juanchich, 2018). The fourth is a belief-bias question (Baron et al., 2015; Markovits & Nantel, 1989), and the fifth is a conjunction-fallacy item (Kahneman & Tversky, 1973; Yilmaz & Saribay, 2017). Reflection scores equal the total number of correct responses, while intuition scores equal the number of intuitive responses.

2.2.3. 4-Component thinking styles questionnaire (4-CTSQ)

The 4-CTSQ was developed by Newton et al. (2024) to assess the multidimensionality of thinking styles through a self-report questionnaire. The scale consists of four subscales, each containing six items: actively open-minded thinking (α = .92), close-minded thinking (α = .87), preference for intuitive thinking (α = .95), and preference for effortful thinking (α = .90). Participants responded to items using a 6-point scale, ranging from 1 (strongly disagree) to 6 (strongly agree).

2.2.4. Raven's progressive matrices

Raven's Standard Progressive Matrices (Raven, 2000) were used to measure reasoning ability. Although Raven's matrices are standardly interpreted as a measure of cognitive ability rather than reflection, the items contain intuitively appealing lures, and performance depends in part on the willingness to spend time to get the right answer, a feature they share with reflection measures (Baron et al., 2015). We therefore treated Raven's scores as an ability-based, convergent indicator of reflective cognition rather than as a direct measure of reflection. Participants completed three 3 × 3 figure matrices with one missing element, selecting the option that best completed the pattern. Scores reflected the total number of correct responses, following the procedure of Mani et al. (2013, Study 2).

2.2.5. Religiosity

Religiosity was measured using a single-item question: 'How religious do you consider yourself'? rated from 0 (not at all religious) to 10 (very religious).

2.2.6. The monetary moral dilemma

To assess utilitarian behavior in a behavioral context, we employed a monetary moral dilemma[3] (Bostyn et al., 2019). Participants were first grouped with six other passive participants, and a monetary arrangement was presented at the beginning of the task (see current arrangement in Figure 1). In the initial arrangement, the participant and one other group member (C1) were set to receive a $1 bonus, while the remaining five participants (C2–C6) received nothing. Participants were then given the option to change the final bonus distribution for others in their group by asking, 'Would you like to change to the alternative arrangement where C1 does not receive any bonus, and instead, five other participants receive $1 bonus each'? Participants were also informed that their decision would have no impact on their own bonus. As in the traditional trolley dilemma, participants chose between a deontological option of not intervening (current arrangement) and a utilitarian option that sacrificed one to benefit five (alternative arrangement). Choosing the alternative arrangement was coded as utilitarian behavior. In line with Bostyn et al.'s (2019) finding that framing (gain vs. loss) does not affect outcomes, we

[3] Although structurally similar to the trolley dilemma in its one-versus-five trade-off, the task in the current study does not capture the personal/impersonal distinction. We therefore refer to it as the monetary moral dilemma to better reflect its incentivized and physically harmless nature.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

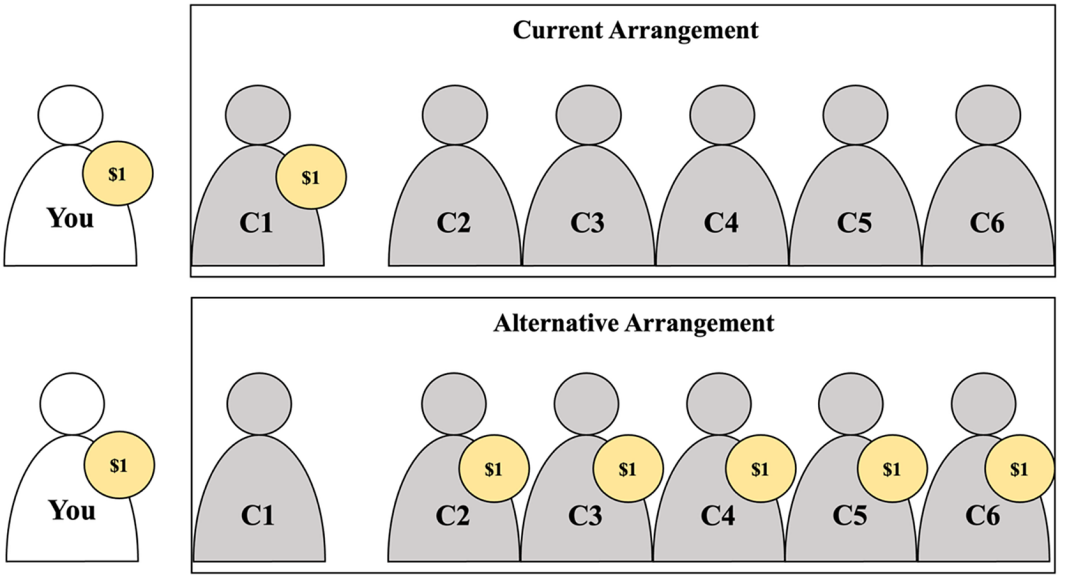


Figure 1. *Bonus distributions for the current and alternative arrangements in the monetary moral dilemma.*

Table 1. *Decision justification questions.*

Justifications	Items
Amoral	Moral reasons did not play an important role in my judgment.
Virtue-ethical	Someone who intentionally harms an innocent person cannot be moral.
Deontological	Intentionally harming an innocent person is against fundamental moral rules and is thus unacceptable regardless of its intended consequence.
Utilitarian	Moral action is what ensures the well-being of maximum number of people.
Fatalistic	It is wrong to interfere with consequences that arise as a result of the natural course of events, no matter what the ensuing harm is.
Inequality aversion	It is wrong to commit actions that increase the inequality of income among people.

used only the loss frame. We also included real monetary incentives, avoided deception, and provided simplified illustrated instructions.

2.2.7. Decision justification questions

After the monetary moral dilemma, participants rated the extent to which each item played a role in their decision-making (i.e., whether to keep the current arrangement or accept the alternative arrangement) on single-item Likert scales (1 = completely disagree, 5 = completely agree) to measure their moral justifications, covering amorality, virtue ethics, deontology, classical utilitarianism, fatalism (Aktas et al., 2017), and inequality aversion to capture fairness concerns (Table 1).

3. Results

3.1 Confirmatory tests

As preregistered, point-biserial correlations were estimated to test the correlations between utilitarian behavior and OUS dimensions of impartial beneficence (H1) and tolerance of instrumental harm (H2).

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

The utilitarian behavior was not significantly associated with either impartial beneficence (r_{pb} = −.08, p = .13) or tolerance of instrumental harm (r_{pb} = −.05, p = .35). To test the link between cognitive reflection and utilitarian behavior in the monetary moral dilemma (H3), a preregistered logistic regression was conducted with CPT reflection scores as the predictor and utilitarian behavior as the outcome. Of 321 participants, 222 chose the utilitarian option. While the model yielded 69.2% classification accuracy, matching the sample's intercept-only base rate, a likelihood ratio test confirmed that adding CPT significantly improved the model's overall fit ($\chi^2(1) = 5.77$, p = .016). Continuous metrics indicated that CPT exerts a statistically meaningful, albeit subtle, effect on the probability of making a utilitarian choice, performing better than random guessing (McFadden's Pseudo-R^2 = .0145, AUC = .585). Specifically, higher CPT reflection scores increased the odds of utilitarian choices (*OR* = 1.19, 95% CI [1.03, 1.38]).[4]

3.2. Exploratory tests

Pearson correlations among utilitarian behavior, OUS dimensions, and cognitive measures are reported in Table 2. Both impartial beneficence and instrumental harm were generally negatively correlated with reflection and positively correlated with intuition. Specifically, impartial beneficence was negatively associated with CPT reflection (r = −.16, p < .001), preference for effortful thinking (r = −.08, p = .008) and Raven's matrices (r = −.08, p = .007), but positively with preference for intuitive thinking (r = .10, p < .001). With a similar pattern, instrumental harm was negatively correlated with actively open-minded thinking (r = −.11, p < .001) and preference for effortful thinking (r = −.13, p < .001) but positively correlated with close-minded thinking (r = .10, p = .001) and preference for intuitive thinking (r = .14, p < .001). In line with our preregistered hypothesis, actual utilitarian decisions 17 months later correlated positively with CPT reflection (r = .13, p = .016), actively open-minded thinking (r = .21, p < .001), and Raven's matrices (r = .26, p < .001), but negatively correlated with close-minded thinking (r = −.12, p = .035) and preference for intuitive thinking (r = −.14, p = .014). Given all reflection measures positively correlated with utilitarian behavior, a multivariable logistic regression was conducted to examine the unique contributions of CPT reflection, cognitive ability (Raven's Matrices), and actively open-minded thinking to utilitarian behavior in the monetary moral dilemma. The model indicated that CPT reflection scores did not significantly predict utilitarian behavior, *OR* = 0.96, 95% CI [0.81, 1.14], p = .653. In contrast, Raven's matrices significantly predicted utilitarian behavior, *OR* = 1.73, 95% CI [1.33, 2.27], p < .001, as did actively open-minded thinking, *OR* = 1.41, 95% CI [1.14, 1.75], p = .001. These results suggest that, when accounting for cognitive ability and thinking style, CPT reflection scores alone did not uniquely predict utilitarian choice, whereas both cognitive ability and actively open-minded thinking were independently associated with greater odds of choosing the utilitarian option.

As preregistered, we also examined whether utilitarian behavior was more strongly associated with impartial beneficence or tolerance of instrumental harm. Although statistical procedures exist for comparing dependent correlations, such comparisons are difficult to interpret because they do not account for potential differences in measurement reliability. We therefore simply note that the two correlations were similar in magnitude and both close to zero.

Analysis of the decision justifications as a validation check of the monetary moral dilemma revealed that only utilitarian justification was positively correlated with utilitarian behavior (r = .24, p < .001), as can be expected, whereas amorality (r = −.20, p < .001), virtue ethics (r = −.24, p < .001), deontology (r = −.28, p < .001), and fatalism (r = −.37, p < .001) were all negatively correlated with utilitarian behavior (Table 3). We also examined the relationship between decision justifications and OUS dimensions (Table 3). Impartial beneficence was negatively correlated with amorality (r = −.13, p = .018) but positively correlated with competing normative frameworks such as virtue ethics (r = .14, p = .014), deontology (r = .15, p = .008), utilitarianism (r = .16, p = .004), and inequality aversion

[4]The point-biserial correlation between CPT reflection scores and utilitarian behavior was also significant, r = .13, p = .016.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

Table 2. *Summary of correlations, means, standard deviations.*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Utilitarian behavior (monetary moral dilemma)	–	–									
2. Instrumental harm	2.97	1.27	−.05 [−.16, .06]								
3. Impartial beneficence	3.49	1.24	−.08 [−.19, .03]	.15** [.09, .20]							
4. CPT reflection score	2.28	1.61	.13* [.02, .24]	−.03 [−.09, .02]	−.16** [−.21, −.10]						
5. Actively open-minded thinking	4.33	1.23	.21** [.10, .31]	−.11** [−.17, −.06]	−.03 [−.09, .03]	.31** [.26, .36]					
6. Close-minded thinking	2.81	1.24	−.12* [−.22, −.01]	.10** [.04, .15]	.04 [−.02, .09]	−.19** [−.25, −.14]	−.40** [−.44, −.35]				
7. Preference for intuitive thinking	3.44	1.22	−.14* [−.24, −.03]	.14** [.08, .19]	.10** [.05, .16]	−.30** [−.35, −.24]	−.47** [−.51, −.43]	.25** [.20, .30]			
8. Preference for effortful thinking	4.56	1.16	.05 [−.06, .15]	−.13** [−.19, −.08]	−.08** [−.13, −.02]	.21** [.15, .26]	.28** [.23, .33]	−.21** [−.27, −.16]	−.20** [−.25, −.15]		
9. Raven's matrices	1.25	1.11	.26** [.15, .36]	−.02 [−.08, .03]	−.08** [−.13, −.02]	.37** [.32, .42]	.21** [.15, .26]	−.15** [−.21, −.10]	−.18** [−.24, −.13]	.16** [.10, .21]	
10. CPT intuition score	2.54	1.54	−.15** [−.26, −.04]	.03 [−.03, .09]	.17** [.11, .22]	−.96** [−.97, −.96]	−.32** [−.37, −.26]	.19** [.14, .25]	.29** [.24, .34]	−.20** [−.26, −.15]	−.35** [−.40, −.30]

Note: Values in square brackets indicate the 95% confidence interval for each correlation. *indicates $p < .05$. **indicates $p < .01$. Sample size is 321 for correlations with the utilitarian behavior in the monetary moral dilemma and 1,200 for others.

Table 3. *Means, standard deviations, and correlations among moral justifications, utilitarian behavior, and OUS dimensions*.

Justifications	*M*	*SD*	Utilitarian behavior	Impartial beneficence	Instrumental harm
Utilitarian	3.69	1.15	.24**	.16**	−.08
			[.14, .34]	[.05, .27]	[−.19, .03]
Deontological	3.30	1.34	−.28**	.15**	−.08
			[−.38, −.18]	[.04, .25]	[−.18, .03]
Virtue-ethical	3.18	1.33	−.24**	.14*	−.12*
			[−.34, −.13]	[.03, .24]	[−.23, −.01]
Amoral	2.55	1.33	−.20**	−.13*	.05
			[−.30, −.09]	[−.24, −.02]	[−.06, .16]
Fatalistic	2.38	1.22	−.37**	.09	−.01
			[−.46, −.27]	[−.02, .20]	[−.12, .10]
Inequality aversion	3.27	1.32	.06	.22**	−.08
			[−.05, .17]	[.11, .32]	[−.18, .03]

Note: Values in square brackets indicate the 95% confidence interval for each correlation. *indicates $p < .05$. **indicates $p < .01$. $N = 321$. Table generated using the apaTables package in R (Stanley, 2021).

($r = .22$, $p < .001$), suggesting a response bias. Meanwhile, instrumental harm was only negatively associated with virtue ethics ($r = -.12$, $p = .030$).

Table 4 presents the correlations between decision justifications of the monetary moral dilemma and reflection measures. Although CPT reflection, actively open-minded thinking, and Raven's matrices predicted utilitarian behavior (see Table 2), none were significantly correlated with utilitarian justification ($r = .01$, $r = .03$, $r = .04$, respectively, and all ps $> .441$). Only close-minded thinking was negatively associated with utilitarian justification ($r = -.14$; $p = .015$). Regarding other moral justifications, CPT reflection was negatively correlated with deontological ($r = -.24$, $p < .001$), virtue-ethical ($r = -.26$, $p < .001$), fatalistic justifications ($r = -.33$, $p < .001$), and inequality aversion ($r = -.17$, $p = .002$). Actively open-minded thinking was also negatively related to deontological ($r = -.19$, $p < .001$), virtue-ethical ($r = -.23$, $p < .001$), and fatalistic justifications ($r = -.34$, $p < .001$), whereas close-minded thinking showed positive correlations with these three ($r = .17$, $p = .002$; $r = .17$, $p = .002$; $r = .12$, $p = .03$) and with amorality ($r = .15$, $p = .005$). Preference for intuitive thinking was positively associated with virtue-ethical ($r = .20$, $p < .001$) and fatalistic justifications ($r = .15$, $p = .009$), while preference for effortful thinking was negatively related only to fatalistic justification ($r = .-12$, $p = .028$). Raven's matrices scores were negatively correlated with deontological ($r = -.14$, $p = .013$), virtue-ethical ($r = -.15$, $p = .007$), and fatalistic justifications ($r = -.23$, $p < .001$). Overall, reflection-related measures (CPT reflection, actively open-minded thinking, and Raven's matrices) were negatively associated with deontological, virtue-ethical, and fatalistic justifications.

To explore the role of religiosity, we first examined bivariate correlations. Religiosity was negatively correlated with utilitarian behavior ($r = -.17$, $p = .003$), positively correlated with impartial beneficence ($r = .09$, $p = .001$), and showed no significant association with instrumental harm ($r = .03$, $p = .261$). Given the significant correlations of religiosity with utilitarian behavior and impartial beneficence, we conducted exploratory moderation analyses. The logistic regression model including CPT reflection, religiosity, and their interaction significantly predicted utilitarian behavior in the monetary moral dilemma, $\chi^2(3) = 17.64$, $p < .001$ (Figure 2). Higher reflection scores increased the odds of utilitarian behavior (*OR* = 1.36, 95% CI, [1.105, 1.676], $p = .004$). While religiosity had no significant direct effect (*OR* = 1.02, 95% CI, [0.916, 1.149], $p = .663$), a significant interaction between reflection and religiosity was observed (*OR* = .95, 95% CI, [0.914, 0.992], $p = .021$), suggesting that the association between reflective thinking and utilitarian behavior weakens as religiosity increases. In contrast, the moderation analysis for impartial beneficence yielded no significant interaction, suggesting that the

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

Table 4. *Correlations among moral justifications in the monetary moral dilemma and reflection measures*

	CPT reflection score	Actively open-minded thinking	Close-minded thinking	Preference for intuitive thinking	Preference for effortful thinking	Raven's matrices
Utilitarian	.01	.03	−.14*	.01	.01	.04
	[−.10, .12]	[−.08, .14]	[−.24, −.03]	[−.10, .12]	[−.10, .12]	[−.07, .15]
Deontological	−.24**	−.19**	.17**	.11	−.03	−.14*
	[−.34, −.13]	[−.29, −.08]	[.06, .28]	[−.00, .21]	[−.14, .08]	[−.24, −.03]
Virtue-ethical	−.26**	−.23**	.17**	.20**	−.05	−.15**
	[−.36, −.15]	[−.33, −.12]	[.06, .27]	[.09, .30]	[−.16, .06]	[−.26, −.04]
Amoral	−.03	−.04	.15**	−.00	−.09	−.02
	[−.14, .08]	[−.14, .07]	[.05, .26]	[−.11, .11]	[−.20, .02]	[−.13, .09]
Fatalistic	−.33**	−.34**	.12*	.15**	−.12*	−.23**
	[−.43, −.23]	[−.43, −.24]	[.01, .23]	[.04, .25]	[−.23, −.01]	[−.33, −.13]
Inequality aversion	−.17**	.02	−.07	.07	−.07	−.05
	[−.28, −.07]	[−.09, .13]	[−.18, .04]	[−.04, .17]	[−.17, .04]	[−.16, .06]

Note: Values in square brackets indicate the 95% confidence interval for each correlation. *indicates $p < .05$. **indicates $p < .01$. $N = 321$. Table generated using the apaTables package in R (Stanley, 2021).

interaction between religiosity and CPT reflection is specific to utilitarian behavior but not impartial beneficence, $b = -0.005$, $SE = 0.006$, $p = .441$.

4. Discussion

The present study investigated whether reflective thinking predicts utilitarian moral behavior over time in an incentivized paradigm that also included various cognitive assessments (CPT, 4-CTSQ, Raven matrices) and self-reported utilitarian measures (OUS dimensions and utilitarian justification). Consistent with the dual-process model of moral judgment (Greene et al., 2001, 2004; Greene, 2008), reflection was positively associated with utilitarian behavior in a monetary moral dilemma measured 17 months after the initial cognitive tasks. This finding extends prior work by demonstrating that reflection predicts actual moral behavior in a context with real stakes. Importantly, utilitarian dimensions measured by the OUS (Kahane et al., 2018) did not predict behavior, underscoring a divergence between self-report and behavioral measures of moral judgment (Bostyn et al., 2018; Sheeran & Webb, 2016).

4.1. Reflection and utilitarian behavior

The link between reflection and utilitarianism was found across multiple indicators of reflective cognition in both confirmatory (CPT reflection) and exploratory findings (Raven's and actively open-minded thinking). While the confirmatory result strengthens correlational support for the dual-process model's core claim that reflective processes facilitate utilitarian judgments (Greene, 2008; Greene & Haidt, 2002), utilitarian behavior may also be driven not only by the tendency to override intuitive responses, as captured by CPT, but also by broader cognitive resources and the motivation to engage in open and effortful thinking. Raven's matrices, though standardly interpreted as a measure of cognitive ability, also share features with reflection measures: the items contain intuitively appealing lures, and willingness to take time is a major determinant of performance (Baron et al., 2015), so the ability and reflection interpretations are not sharply separable in practice. As these are exploratory, they should be interpreted with caution, but they point to a theoretically meaningful distinction between cognitive

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

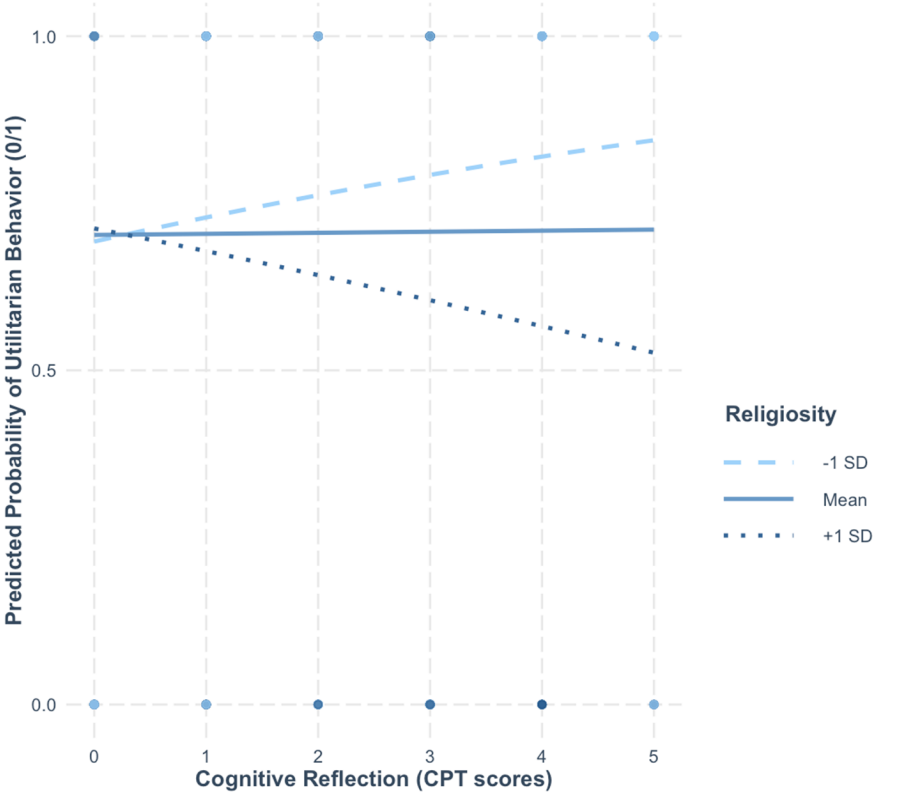


Figure 2. *Interaction plot of religiosity and CPT reflection score on predicted probability of utilitarian behavior.*

reflection, ability, and motivation as separable contributors to utilitarian behavior, warranting further investigation.

Unlike sacrificial dilemmas—often criticized for conflating utilitarian reasoning with antisocial traits such as psychopathy (Bartels & Pizarro, 2011; Djeriouat & Trémolière, 2014; Kahane et al., 2015; Zamora et al., 2022)—the monetary moral task involves no physical harm (Bostyn et al., 2019). At the same time, the monetary dilemma does not involve using one person as a means to benefit others in the way footbridge dilemmas do. Thus, the emotional processes posited by dual-process models to underlie deontological judgments may not be engaged to the same extent, as taking money from one person is unlikely to elicit the same degree of emotional intensity as physical harm.

We showed that the utilitarian choice in the monetary dilemma reflects a distinct utilitarian motivation, given that it was positively correlated only with utilitarian justification. It was also not significantly correlated with the inequality-aversion justification, suggesting that fairness concerns, which might otherwise provide non-utilitarian grounds for redistribution, are unlikely to account for the utilitarian choice. In contrast, the non-utilitarian choice correlated with multiple moral justifications, including deontology, fatalism, virtue ethics, and amorality, suggesting a heterogeneous set of moral and non-moral considerations rather than deontological reasoning alone. A similar pattern has been found in real-world moral dilemmas in lab settings, where utilitarian decisions were driven by cost-benefit motivation, whereas non-harming decisions reflected a range of moral motivations (Bostyn et al., 2025). These findings suggest that deontological principles, as traditionally used in the moral dilemma literature, may not fully account for non-harming behavior, especially when behavioral measures are incorporated.

The 17-month interval between the measurement of reflection and the behavioral task further reduces concerns about demand effects, as participants were unlikely to infer the study's hypotheses. This

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

temporal separation suggests that individual differences in reflection are predictive of moral behavior, converging with meta-analytic evidence on reflective and intuitive thinking in moral dilemmas (Capraro, 2024; Fahrenwaldt et al., 2025).

4.2. Limitations of the Oxford Utilitarianism Scale

The OUS (Kahane et al., 2018) was developed to capture utilitarian tendencies by distinguishing between instrumental harm and impartial beneficence. In the present study, however, neither dimension predicted actual behavior in the monetary moral dilemma. Impartial beneficence correlated with a range of moral justifications, including both deontological and utilitarian ones, raising concerns about its construct validity. These findings are consistent with critiques that impartial beneficence may reflect general altruism or moral expansiveness rather than utilitarian reasoning per se (Mihailov, 2023; Simpson & Nash, 2024), and with evidence suggesting that OUS dimensions may reflect reduced sensitivity to moral norms rather than greater sensitivity to outcomes (Körner et al., 2020). By contrast, utilitarian decision in the monetary moral dilemma was positively associated with utilitarian justification and negatively with deontological, virtue-ethical, and fatalistic ones, supporting its validity as a behavioral measure.

Furthermore, exploratory analyses showed that actively open-minded thinking correlated positively with utilitarian behavior, while close-mindedness and intuitive thinking correlated negatively. The OUS dimensions, in contrast, were positively associated with intuition but negatively with reflection. Moreover, dual-process predictions do not consistently map onto OUS dimensions, as reflection manipulations fail to reliably influence either dimension (Capraro et al., 2019). Thus, the OUS may capture dispositional tendencies toward intuitive reasoning rather than the reflective processes emphasized in dual-process models. Given the limitations of OUS, novel conceptualizations of utilitarianism may be needed. For instance, Antoniou et al. (2023) propose a revised two-dimensional model that integrates emotion into Kahane et al.'s (2018) two-dimensional account, providing a more comprehensive understanding of utilitarian reasoning.

Taken together, these results highlight the limitations of relying solely on self-report measures. While self-report scales such as OUS have good psychometric properties, they may capture socially desirable attitudes or have limited conceptual capacity to capture utilitarianism as theorized in dual-process models. Additionally, earlier findings emphasize the value of behavioral tasks for assessing utilitarian decision-making, addressing the external validity concerns of existing moral dilemma measures (Bostyn et al., 2018; Clarkson, 2022; Sheeran & Webb, 2016). Accordingly, this study also showed that incentivized behavioral tasks offer a more ecologically valid measure of utilitarian decision-making.

4.3. The role of religiosity

One noteworthy exploratory finding of this study is that religiosity moderated the relationship between reflection and utilitarian behavior. Although religiosity itself was not directly associated with utilitarian choices, reflection predicted utilitarianism more strongly among non-believers than believers. This moderation aligns with research emphasizing that reflection may not be uniformly truth-oriented but depends on the epistemic norms individuals endorse (Baron, 2020; Saribay & Yilmaz, 2024; Yilmaz, 2021).

For non-believers, reflection is guided by evidence-based reasoning and belief updating (Bahçekapili & Yilmaz, 2017; Metz et al., 2018; Yilmaz & Isler, 2019), promoting impartial deliberation and outcome-based judgments. For believers, who often treat divine authority as a source of moral truth (Baron et al., 2015; Piazza & Landy, 2013; Simpson et al., 2016), reflection may instead reinforce convictions grounded in sacred texts. Our findings thus extend the literature indicating negative correlations between religiosity and utilitarian judgments (Saroglou & Craninx, 2021) by demonstrating empirically that religiosity conditions the reflection–utilitarianism link, situating dual-process models within broader sociocultural contexts (Saribay & Yilmaz, 2024; Yilmaz & Isler, 2025).

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

4.4. Limitations and future directions

Several limitations should be acknowledged. Although assessing the predictor and outcome variables in the same participants over a 17-month interval adds temporal depth and allows for examination of long-term associations, the analyses were correlational and do not establish causality. Furthermore, because different constructs were measured at Time 1 and Time 2, the design does not reduce within-subject variability in the same way as repeated measures. Although the temporal stability of cognitive reflection (Stagnaro et al., 2018) and sensitivity to consequences in moral dilemmas were found (Luke & Gawronski, 2022), uncontrollable life events (e.g., elections, severe illness, economic hardship) during long periods, similar to the current study, may have influenced participants' moral decisions. Future research should employ experimental[5] or longitudinal designs that can more directly test causal mechanisms and capture potential changes in moral behavior across longer intervals while retaining the external validity of behavioral tasks.

Second, utilitarian behavior was assessed with a single monetary moral dilemma. While this task improves upon sacrificial dilemmas by removing physical harm and incorporating real incentives without any experimental deception, it captures only one dimension of utilitarian reasoning and thus has limited content validity as a measure of broader utilitarian tendencies. Developing a broader set of incentivized tasks, such as redistribution, impartial aid, or public goods games, would provide a more comprehensive picture of utilitarian moral psychology.

Third, moral justifications (utilitarianism, deontology, virtue ethics, amorality, fatalism, and inequality aversion) were measured with single items, limiting psychometric robustness and the depth of insights into participants' normative views. Multi-item scales or qualitative methods could enrich understanding of the cognitive and normative frameworks people invoke. Similarly, religiosity was assessed with a single self-report item; because different operationalizations of religiosity can relate differently to the two utilitarian dimensions (Paruzel-Czachura & Charzyńska, 2023), future work should employ multidimensional measures of religiosity. Finally, the sample was drawn from the U.S. and may not generalize across cultures. Cross-cultural research is especially important given differences in religiosity, epistemic norms, and moral priorities, and could test whether the moderating role of religiosity is consistent across contexts.

Lastly, while the current study is grounded in Greene's original dual-process model for understanding the role of reflective cognition in utilitarian decision-making, this framework has also been subject to criticism and refinement. Specifically, Bago and De Neys (2019) found that intuitive responses in moral dilemmas are not exclusively deontological and that utilitarian judgments do not necessarily require deliberative override of intuitive responses. Future research should also examine the precise cognitive mechanisms underlying utilitarian behavior using behavioral paradigms that can more directly test the timing and nature of intuitive and deliberative processes.

5. Conclusion

This preregistered study provides evidence that reflection predicts utilitarian behavior in an incentivized economic task, whereas self-reported OUS dimensions do not. These findings extend the dual-process model of moral judgment by moving beyond hypothetical dilemmas to actual behavior, while also cautioning against overreliance on self-report scales such as the OUS. In addition, the moderating role of religiosity is a result worth following up, as reflection's impact on moral decision-making may depend on epistemic norms that guide reasoning.

By integrating cognitive, self-report, and behavioral measures across time, this study advances understanding of the reflection–utilitarianism link and suggests dual-process models within broader

[5] A truly impulsive participant might respond randomly, and since the utilitarian option requires reflective thinking, impulsiveness might be a potential third variable worth examining. Results on response times (time spent until decision page submission) on the task showed that choosing the utilitarian option took longer to respond than the non-utilitarian option, $t(136.98) = 3.06$, $p = .003$; log-transformed means: 4.01 versus 3.67, suggesting that impulsiveness is unlikely to account for the findings. In addition, in a separate study (N =1547) using a continuous version of the task, utilitarian behavior was also not significantly correlated with an 'I decided randomly' item ($r = -.06$, $p < .001$), which was given alongside justification items.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

sociocultural contexts. In doing so, it contributes to a more ecologically grounded and nuanced account of individual differences in moral cognition.

Acknowledgments. Parts of this research were included in the first author's master's thesis (R.V.) submitted to Kadir Has University and conducted under the last author's supervision (O.Y.). We are grateful to Mariola Paruzel-Czachura, Jonathon Baron, and an anonymous reviewer for their constructive and valuable feedback during the revision process. We used ChatGPT (GPT-4o; chat.openai.com) to proofread and improve language of our text. ChatGPT revised existing content and did not generate any part of the article independently.

Data availability statement. The data, analysis code, codebook, and materials of this study are openly available in the Open Science Framework (OSF) at https://osf.io/6r7gn/.

Funding statement. This research was supported by the Templeton Religion Trust [grant number: TRT0424]. The first author was supported by the 2210-A Program of The Research and Technological Council of Turkey (TUBITAK).

Competing interest. The authors have no competing interests to declare that are relevant to the content of this article.

Ethical standards. We received ethical approval from the Human Research Ethics Committee of Kadir Has University (Report number: E-82741295-604.01.01-29719).

References

Aktas, B., Yilmaz, O., & Bahçekapili, H. G. (2017). Moral pluralism on the trolley tracks: Different normative principles are used for different reasons in justifying moral judgments. *Judgment and Decision Making*, *12*(3), 297–307. https://doi.org/10.1017/S1930297500005891

Antonenko Young, O., Willer, R., & Keltner, D. (2013). "Thou shalt not kill": Religious fundamentalism, conservatism, and rule-based moral processing. *Psychology of Religion and Spirituality*, *5*(2), 110. https://doi.org/10.1037/a0032262

Antoniou, R., Romero-Kornblum, H., Young, J. C., You, M., Kramer, J. H., Rankin, K. P., & Chiong, W. (2023). Contrasting two models of utilitarian reasoning. *Heliyon*, *9*(7), e17498. https://doi.org/10.1016/j.heliyon.2023.e17498

Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General*, *148*(10), 1782–1801. https://doi.org/10.1037/xge0000533

Bahçekapili, H. G., & Yilmaz, O. (2017). The relation between different types of religiosity and analytic cognitive style. *Personality and Individual Differences*, *117*, 267–272. https://doi.org/10.1016/j.paid.2017.06.013

Baron, J. (2020). Religion, cognitive style, and rational thinking. *Current Opinion in Behavioral Sciences*, *34*, 64–68. https://doi.org/10.1016/j.cobeha.2019.12.015

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the cognitive reflection test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, *4*(3), 265–284. https://doi.org/10.1016/j.jarmac.2014.09.003

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*(1), 154–161. https://doi.org/10.1016/j.cognition.2011.05.010

Bates, T. C. (2025). Genetic origins of Utilitarian versus Kantian moral philosophy in heritable motivations for egalitarian beneficence and coercive redistribution. *Personality and Individual Differences*, *238*, 113109. https://doi.org/10.1016/j.paid.2025.113109

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, *8*(9), 536–554. https://doi.org/10.1111/spc3.12131

Bostyn, D. H., Gouwy, M. C., De Craene, E., Vanmechelen, C., Scheirlinckx, J., Tissot, T. T., . . . Roets, A. (2025). Beyond hypothetical trolleys: Moral choices and motivations in a real-life sacrificial dilemma. *Journal of Personality and Social Psychology*, *129*(5), 834–849. https://doi.org/10.1037/pspa0000463

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science*, *29*(7), 1084–1093. https://doi.org/10.1177/0956797617752640

Bostyn, D. H., Sevenhant, S., & Roets, A. (2019). Beyond physical harm: How preference for consequentialism and primary psychopathy relate to decisions on a monetary trolley dilemma. *Thinking & Reasoning*, *25*(2), 192–206. https://doi.org/10.1080/13546783.2018.1497536

Capraro, V. (2024). The dual-process approach to human sociality: Meta-analytic evidence for a theory of internalized heuristics for self-preservation. *Journal of Personality and Social Psychology*. Advance online publication. https://doi.org/10.1037/pspa0000375

Capraro, V., Everett, J. A. C., & Earp, B. D. (2019). Priming intuition disfavors instrumental harm but not impartial beneficence. *Journal of Experimental Social Psychology*, *83*, 142–149. https://doi.org/10.1016/j.jesp.2019.04.006

Clarkson, E. (2022). Moral behavior in games: A review and call for additional research. *New Ideas in Psychology*, *64*, 100912. https://doi.org/10.1016/j.newideapsych.2021.100912

Côté, S., Piff, P. K., & Willer, R. (2013). For whom do the ends justify the means? Social class and utilitarian moral judgment. *Journal of Personality and Social Psychology*, *104*(3), 490. https://doi.org/10.1037/a0030931

Crone, D. L., & Laham, S. M. (2017). Utilitarian preferences or action preferences? De-confounding action and moral code in sacrificial dilemmas. *Personality and Individual Differences*, *104*, 476–481. https://doi.org/10.1016/j.paid.2016.09.022

Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience*, *7*(3), 269–279. https://doi.org/10.1080/17470919.2011.614000

Djeriouat, H., & Trémolière, B. (2014). The Dark Triad of personality and utilitarian moral judgment: The mediating role of Honesty/Humility and Harm/Care. *Personality and Individual Differences*, *67*, 11–16. https://doi.org/10.1016/j.paid.2013.12.026

Fahrenwaldt, A., Olsen, J., Rahal, R.-M., & Fiedler, S. (2025). Intuitive deontology? A systematic review and multivariate, multilevel meta-analysis of experimental studies on the psychological drivers of moral judgments. *Psychological Bulletin*, *151*(4), 428–454. https://doi.org/10.1037/bul0000472

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Gold, N., Pulford, B. D., & Colman, A. M. (2015). Do as I say, don't do as I do: Differences in moral judgments do not translate into differences in decisions in real-life trolley problems. *Journal of Economic Psychology*, *47*, 50–61. https://doi.org/10.1016/j.joep.2015.01.001

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523. https://doi.org/10.1016/S1364-6613(02)02011-9

Greene, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 3. The neuroscience of morality: Emotion, brain disorders, and development* (pp. 35–80). Boston Review.

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*(2), 389–400. https://doi.org/10.1016/j.neuron.2004.09.027

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Isler, O., & Yilmaz, O. (2023). How to activate intuitive and reflective thinking in behavior research? A comprehensive examination of experimental techniques. *Behavior Research Methods*, *55*, 3679–3698. https://doi.org/10.3758/s13428-022-01984-4

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. https://doi.org/10.1016/j.cognition.2014.10.005

Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131–164. https://doi.org/10.1037/rev0000093

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251. https://doi.org/10.1037/h0034747

Körner, A., & Deutsch, R. (2022). Deontology and utilitarianism in real life: A set of moral dilemmas based on historic events. *Personality and Social Psychology Bulletin*, *49*(10), 1511–1528. https://doi.org/10.1177/01461672221103058

Körner, A., Deutsch, R., & Gawronski, B. (2020). Using the CNI Model to investigate individual differences in moral dilemma judgments. *Personality and Social Psychology Bulletin*, *46*(9), 1392–1407. https://doi.org/10.1177/0146167220907203

Luke, D. M., & Gawronski, B. (2022). Temporal stability of moral dilemma judgments: A longitudinal analysis using the CNI model. *Personality and Social Psychology Bulletin*, *48*(8), 1191–1203.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science*, *341*(6149), 976–980. https://doi.org/10.1126/science.1238041

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, *17*(1), 11–17. https://doi.org/10.3758/bf03199552

McPhetres, J., Conway, P., Hughes, J. S., & Zuckerman, M. (2018). Reflecting on God's will: Reflective processing contributes to religious peoples' deontological dilemma responses. *Journal of Experimental Social Psychology*, *79*, 301–314.

Metz, S. E., Weisberg, D. S., & Weisberg, M. (2018). Non-scientific criteria for belief sustain counter-scientific beliefs. *Cognitive Science*, *42*(5), 1477–1503. https://doi.org/10.1111/cogs.12584

Mihailov, E. (2023). Measuring impartial beneficence: A Kantian perspective on the Oxford Utilitarianism Scale. *Review of Philosophy and Psychology*, *14*, 989–1004. https://doi.org/10.1007/s13164-021-00600-2

Newton, C., Feeney, J., & Pennycook, G. (2024). On the disposition to think analytically: Four distinct intuitive-analytic thinking styles. *Personality and Social Psychology Bulletin*, *50*(6), 906–923. https://doi.org/10.1177/01461672231154886

Ni, B. K., Burns, B. D., Mak, K. K., Lah, S., Silva, D. S., Goldwater, M. B., & Kleitman, S. (2023). To kill or not to kill: A systematic literature review of high-stakes moral decision-making measures and their psychometric properties. *Frontiers in Psychology*, *13*, 1063607. https://doi.org/10.3389/fpsyg.2022.1063607

Oshiro, B., McAuliffe, W. H., Santos, A. C., Findor, A., Kuzminska, A. O., & Flake, J. K. (2024). Structural validity evidence for the Oxford Utilitarianism Scale across 15 languages. *Psychological Test Adaptation and Development*, *5*(1), 175–191. https://doi.org/10.1027/2698-1866/a000061

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

Paruzel-Czachura, M., & Farny, Z. (2023). Psychopathic traits and utilitarian moral judgment revisited. *Personality and Social Psychology Bulletin*, *50*(9), 1368–1385. https://doi.org/10.1177/01461672231169105

Patil, I., Zucchelli, M. M., Kool, W., Campbell, S., Fornasier, F., Calò, M., . . . Cushman, F. (2021). Reasoning supports utilitarian resolutions to moral dilemmas across diverse measures. *Journal of Personality and Social Psychology*, *120*(2), 443–460. https://doi.org/10.1037/pspp0000281

Piazza, J., & Landy, J. F. (2013). 'Lean not on your own understanding': Belief that morality is founded on divine authority and non-utilitarian moral judgments. *Judgment and Decision Making*, *8*(6), 639–661. https://doi.org/10.1017/S1930297500004691

Raven, J. (2000). The Raven's progressive matrices: change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48. https://doi.org/10.1006/cogp.1999.0735

Rehren, P. (2024). The effect of cognitive load, ego depletion, induction and time restriction on moral judgments about sacrificial dilemmas: A meta-analysis. *Frontiers in Psychology*, *15*, 1388966. https://doi.org/10.3389/fpsyg.2024.1388966

Saribay, S. A., & Yilmaz, O. (2024). *Reflection and intuition in a crisis-ridden world: Thinking hard or hardly thinking?*. Taylor & Francis.

Saroglou, V., & Craninx, M. (2021). Religious moral righteousness over care: A review and a meta-analysis. *Current Opinion in Psychology*, *40*, 79–85. https://doi.org/10.1016/j.copsyc.2020.09.002

Sheeran, P., & Webb, T. L. (2016). The intention–behavior gap. *Social and Personality Psychology Compass*, *10*(9), 503–518. https://doi.org/10.1111/spc3.12265

Simpson, A., Piazza, J., & Rios, K. (2016). Belief in divine moral authority: Validation of a shortened scale with implications for social attitudes and moral cognition. *Personality and Individual Differences*, *94*, 256–265. https://doi.org/10.1016/j.paid.2016.01.032

Simpson, D., & Nash, K. (2024). Applying process dissociation to self-sacrificial moral dilemmas: Extending the dual-process model. *Cognition*, *250*, 105873. https://doi.org/10.1016/j.cognition.2024.105873

Sirota, M., & Juanchich, M. (2018). Effect of response format on cognitive reflection: Validating a two- and four-option multiple choice question version of the cognitive reflection test. *Behavior Research Methods*, *50*(6), 2511–2522. https://doi.org/10.3758/s13428-018-1029-4

Stagnaro, M. N., Pennycook, G., & Rand, D. G. (2018). Performance on the Cognitive Reflection Test is stable across time. *Judgment and Decision Making*, *13*(3), 260–267.

Stanley, D. J. (2021). apaTables: Create American Psychological Association (APA) *Style* Tables. R package version 2.0.8. https://CRAN.R-project.org/package=apaTables

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*(1), 72–81. https://doi.org/10.1177/1745691614556679

Yilmaz, O. (2021). Cognitive styles and religion. *Current Opinion in Psychology*, *40*, 150–154. https://doi.org/10.1016/j.copsyc.2020.09.014

Yilmaz, O., & Isler, O. (2019). Reflection increases belief in God through self-questioning among non-believers. *Judgment and Decision Making*, *14*(6), 649–657. https://doi.org/10.1017/S1930297500005374

Yilmaz, O., & Isler, O. (2025). Epistemic norm differences matter. *Religion, Brain & Behavior*, 1–5. https://doi.org/10.1080/2153599X.2025.2507657

Yilmaz, O., & Saribay, S. A. (2017). The relationship between cognitive style and political orientation depends on the measures used. *Judgment and Decision Making*, *12*(2), 140–147. https://doi.org/10.1017/S1930297500005684

Zamora, J., Ungson, N. D., & Seidman, G. (2022). The end justifies the me: Self-interest moderates the relationship between Dark Triad traits and utilitarian moral decisions. *Personality and Individual Differences*, *184*, 111134. https://doi.org/10.1016/j.paid.2021.111134

Appendix

Oxford Utilitarianism Scale (OUS)

The following statements focus on your thoughts in different situations. Please indicate how much you agree or disagree with each of these statements using the scale provided below. (1 = Strongly Disagree; 7 = Strongly Agree)

1) If the only way to save another person's life during an emergency is to sacrifice one's own leg, then one is morally required to make this sacrifice.

2) From a moral point of view, we should feel obliged to give one of our kidneys to a person with kidney failure since we do not need two kidneys to survive, but really only one to be healthy.

3) From a moral perspective, people should care about the well-being of all human beings on the planet equally; they should not favor the well-being of people who are especially close to them either physically or emotionally.

4) It is just as wrong to fail to help someone as it is to actively harm them yourself.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

5) It is morally wrong to keep money that one does not really need if one can donate it to causes that provide effective help to those who will benefit a great deal.

6) It is morally right to harm an innocent person if harming them is a necessary means to helping several other innocent people.

7) If the only way to ensure the overall well-being and happiness of the people is through the use of political oppression for a short, limited period, then political oppression should be used.

8) It is permissible to torture an innocent person if this would be necessary to provide information to prevent a bomb going off that would kill hundreds of people.

9) Sometimes it is morally necessary for innocent people to die as collateral damage—if more people are saved overall.

4-Component Thinking Style Questionnaire (4-CTSQ)

Please indicate to what extent you agree or disagree with the following statements. (1 = Strongly Disagree; 6 = Strongly Agree)

1) It is important to be loyal to your beliefs even when evidence is brought to bear against them. (R)

2) Whether something feels true is more important than evidence. (R)

3) Just because evidence conflicts with my current beliefs does not mean my beliefs are wrong. (R)

4) There may be evidence that goes against what you believe, but that does not mean you have to change your beliefs. (R)

5) Even if there is concrete evidence against what you believe to be true, it is OK to maintain cherished beliefs. (R)

6) Regardless of the topic, what you believe to be true is more important than evidence against your beliefs. (R)

7) I think there are many wrong ways, but only one right way, to almost anything.

8) In my experience, the truth is often black and white.

9) Truth is never relative.

10) The truth does not change.

11) Either something is true, or it is false; there is nothing in between.

12) There is no middle ground between what is true and what is false.

13) I like to rely on my intuitive impressions.

14) I believe in trusting my hunches.

15) When I make decisions, I tend to rely on my intuition.

16) Using my 'gut-feelings' usually works well for me in figuring out problems in my life.

17) Intuition is the best guide in making decisions.

18) I often go by my instincts when deciding on a course of action. (18)

19) I am not that good at figuring out complicated problems. (R)

20) Thinking is not my idea of an enjoyable activity. (R)

21) I try to avoid situations that require thinking in depth about something. (R)

22) I am not a very analytical thinker. (R)

23) Reasoning things out carefully is not one of my strong points. (R)

24) Thinking hard and for a long time about something gives me little satisfaction. (R)

Cognitive Performance Test

1) A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost?

◯ 5 pence (1)
◯ 10 pence (2)
◯ 9 pence (3)
◯ 1 pence (4)

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

◯ 5 minutes (1)
◯ 100 minutes (2)
◯ 20 minutes (3)
◯ 500 minutes (4)

3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

◯ 47 days (1)
◯ 24 days (2)
◯ 12 days (3)
◯ 36 days (4)

4) All living things need water. Roses need water. If these two statements are true, can we conclude from them that roses are living things?

◯ Yes (1)
◯ No (2)

5) Claire is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?

◯ Claire is a bank teller (1)
◯ Claire is a bank teller and is active in the feminist movement (2)

Raven's Progressive Matrices

Please try to correctly answer the next three questions. You will receive 5 pence for each correct answer in addition to the participation fee.

On each of the next three screens, you will be shown an image with a missing piece. You need to choose the missing piece that logically completes the image from among the given options. Please see below for a simple example, where the correct answer is '2'.

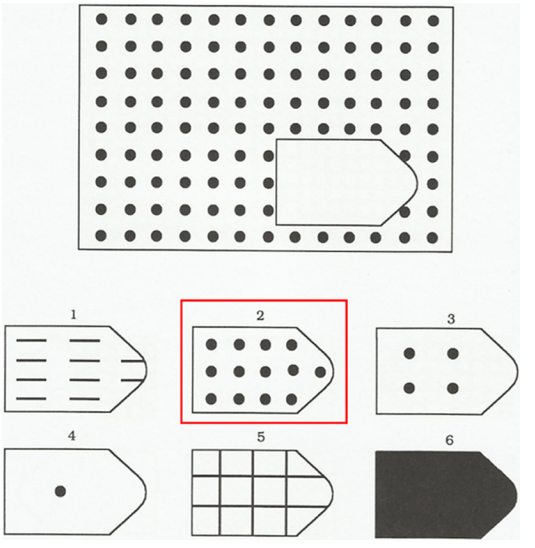


Figure A1.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

1)

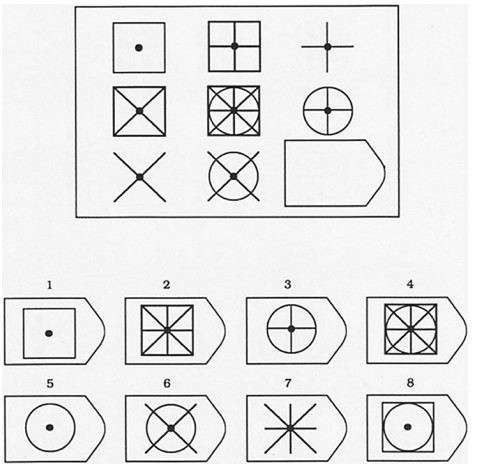


Figure A2.

- ○ 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)

2)

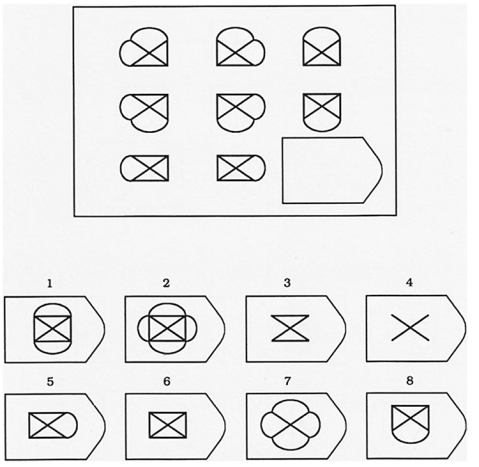


Figure A3.

- ○ 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

3)

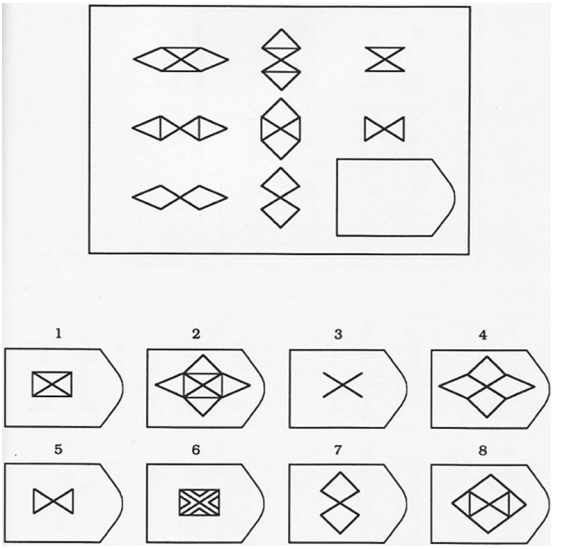


Figure A4.

◯ 1 (1)
◯ 2 (2)
◯ 3 (3)
◯ 4 (4)
◯ 5 (5)
◯ 6 (6)
◯ 7 (7)
◯ 8 (8)

Monetary Moral Dilemma (MMD)

In this task, you will interact with six other participants (Participant C1 to C6).

You are randomly selected to make a decision in this task while the six other participants will be passive. You will receive $1 for this task, regardless of your decision.

One of the six other participants (Participant C1) has also been randomly selected by a computer to receive $1 from this task, whereas the other five were not chosen to receive any bonus. This means that, in the current arrangement, only two are set to receive a bonus out of the seven participants grouped for this task: you and Participant C1.

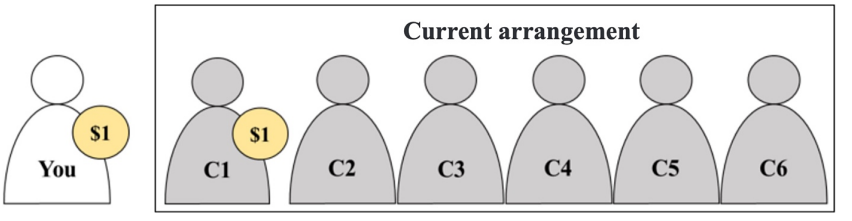


Figure A5.

However, you will determine who will ultimately receive a bonus from this task. As things stand, only you and Participant C1 will receive a $1 bonus, whereas the other five participants will not receive any bonus. Now, you have an opportunity to change this arrangement.

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

If you wish, you can change to an alternative arrangement where Participant C1 does not receive any bonus, and instead, each of the five other participants receive $1 bonus.

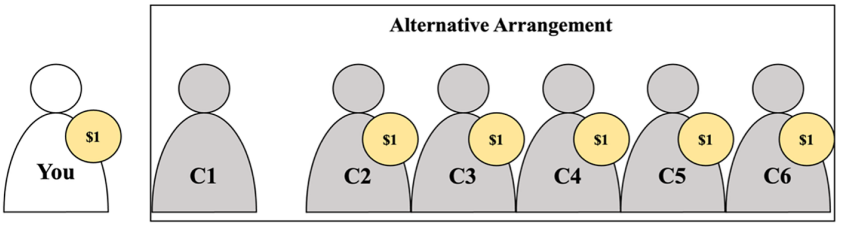


Figure A6.

If you make this change in the arrangements then the one participant who was originally selected to receive a bonus (Participant C1) will not receive it, but all other participants (Participant C2 to C6) will receive bonuses. Whatever you decide will have no impact on your own bonus.

Would you like to change to the alternative arrangement where Participant C1 does not receive any bonus, and instead, five other participants receive $1 bonus each?

Please select an option.

◯ *Change to the alternative arrangement.*
◯ *Do not change the current arrangement.*

Decision Justification in the Monetary Moral Decisions

Please rate to what extent each of the following played a role while making your decision in this task:
1 = Completely disagree; 5 = Completely agree

1) Moral reasons did not play an important role in my judgment.
2) Someone who intentionally harms an innocent person cannot be moral.
3) Intentionally harming an innocent person is against fundamental moral rules and is thus unacceptable regardless of its intended consequence.
4) Moral action is what ensures the well-being of maximum number of people.
5) It is wrong to interfere with consequences that arise as a result of the natural course of events no matter what the ensuing harm is.
6) It is wrong to commit actions that increase the inequality of income among people.

Demographic Form

Gender What is your gender?

◯ Male (1)
◯ Female (2)
◯ Non-binary/third gender (3)
◯ Prefer not to say (4)

Age What is your age in years?

_____

Education level What is the highest level of education you have completed?

◯ Some high school or less (1)
◯ High school diploma or GED (2)
◯ Some college, but no degree (3)
◯ Associates or technical degree (4)
◯ Bachelor's degree (5)
◯ Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, DDS etc.) (6)
◯ Prefer not to say (7)

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

Household income What was your total household income before taxes during the past 12 months?

◯ Less than $25,000 (1)
◯ $25,000–$49,999 (2)
◯ $50,000–$74,999 (3)
◯ $75,000–$99,999 (4)
◯ $100,000–$149,999 (5)
◯ $150,000 or more (6)

Number of people living together What is the number of people you live together (including you)?
Ladder (Subjective SES) Imagine that this ladder pictures how American society is set up.

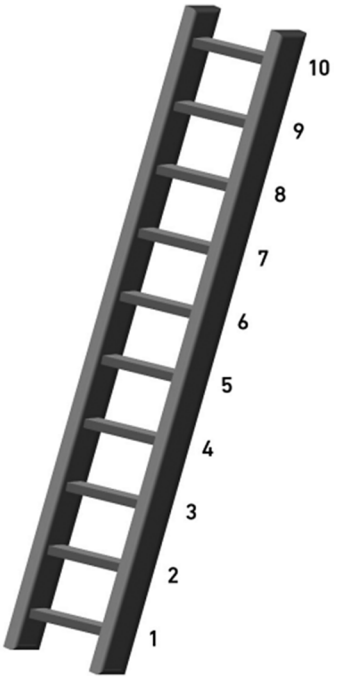


Figure A7.

At the top of the ladder are the people who are the best off—they have the most money, the highest amount of schooling, and the jobs that bring the most respect.

At the bottom are people who are the worst off—they have the least money, little or no education, no job, or jobs that no one wants or respects.

Now think about your family. Please tell us where you think your family would be on this ladder. Select the place that best represents where your family would be on this ladder.

□ 10
□ 9
□ 8
□ 7
□ 6
□ 5
□ 4
□ 3
□ 2
□ 1

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

Belief in God (categorical) Do you believe in God?

◯ Yes (1)
◯ No (2)

Belief in God (continuous) Do you believe in God? (1 = Definitely not; 10 = Definitely yes)

Religious affiliation Please choose the option below that best describes your religious affiliation.
□ Christian, Anglican (1)
□ Christian, Baptist (2)
□ Christian, Catholic (3)
□ Christian, Evangelical (4)
□ Christian, Mormon (5)
□ Christian, Orthodox (6)
□ Christian, Protestant (7)
□ Christian, Unitarian (8)
□ Christian, Other (9)
□ Hindu (10)
□ Jewish, Conservative (11)
□ Jewish, Orthodox (12)
□ Jewish, Reform (13)
□ Jewish, Ultra Orthodox (14)
□ Jewish, Other (15)
□ Pagan (16)
□ Muslim, Kharijite (17)
□ Muslim, Shia (18)
□ Muslim, Sunni (19)
□ Muslim, Other (20)
□ Buddhist (21)
□ Atheist (22)
□ Agnostic (23)
□ Other (24)

Level of religious commitment What is your level of commitment to the religious affiliation you have chosen above? (0 = Extremely Low; 10 = Extremely High)

Religiosity How religious do you describe yourself? (0 = Not at all religious, 10 = Very religious)
General ideology How would you place your political views on this scale, generally speaking? (0 = Extremely Liberal; 10 = Extremely Conservative)

Ideology—social issues When it comes to social issues, how liberal or conservative are you? (0 = Extremely Liberal; 10 = Extremely Conservative)

Ideology—economic issues When it comes to economic issues, how liberal or conservative are you? (0 = Extremely Liberal; 10 = Extremely Conservative)

Ideology (categorical) Where would you place yourself along the political spectrum?

◯ Conservative (1)
◯ Moderate (2)
◯ Liberal (3)
◯ Libertarian (4)
◯ Other (5)

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press

Happiness How happy do you feel in general? (0 = Extremely Unhappy, 10 = Extremely Happy)

Meaning in life Please take a moment to think about what makes your life feel important to you. Please respond to the following statement as truthfully and accurately as you can, and also please remember that it is very subjective question and that there is no right or wrong answer. Please answer according to the scale below: My life has a clear sense of purpose.

◯ Absolutely Untrue (1)
◯ Mostly Untrue (2)
◯ Somewhat Untrue (3)
◯ Cannot Say True or False (4)
◯ Somewhat True (5)
◯ Mostly True (6)
◯ Absolutely True (7)

Belief in evolution Do you believe in evolution? (0 = Definitely not; 10 = Definitely yes)
Belief in karma Do you believe in karma? (0 = Definitely not; 10 = Definitely yes)
Belief in witchcraft Do you believe in witchcraft? (0 = Definitely not; 10 = Definitely yes)
Belief in alternative medicine Do you believe in alternative medicine? (0 = Definitely not; 10 = Definitely yes)
Life satisfaction All things considered, how satisfied are you with your life as a whole these days? (1 = Completely dissatisfied; 10 = Completely satisfied)

Cite this article: Vurgun, R., Acem, E., Seker, F., Dogruyol, B., Bahcekapili, H. G., Isler, O., and Yilmaz, O. (2026). Cognitive reflection predicts utilitarian behavior in the monetary moral dilemma after 17 months. *Judgment and Decision Making*, e30. https://doi.org/10.1017/jdm.2026.10051

https://doi.org/10.1017/jdm.2026.10051 Published online by Cambridge University Press