# I Think, Therefore I Forgive: The Role of Reasoning in Forgiving Accidental Harms

**Indrajeet Patil**
Max Planck Institute for Human Development

**Bastien Trémolière** ( bastien.tremoliere@gmail.com )
University of Nîmes

**Research Article**

**Keywords:** Third-party moral evaluation, Accidental transgressions, Reasoning and Thinking Style, Dual-Process toolbox

**Posted Date:** April 21st, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-379834/v1



**Version of Record:** A version of this preprint was published at Scientific Reports on July 13th, 2021. See the published version at https://doi.org/10.1038/s41598-021-93908-z.

# I think, therefore I forgive: The role of reasoning in forgiving accidental harms

Indrajeet Patil[1], Bastien Trémolière[2]✉

[1]Center for Humans and Machines, Max Planck Institute for Human Development, Berlin, Germany.

[2] APSY-V, University of Nîmes. F-30021 Nîmes Cedex 1, France.

✉Correspondence and requests for materials should be addressed to:
Bastien Trémolière (bastien.tremoliere@gmail.com)

**Abstract**

People experience a strong conflict while evaluating actors who unintentionally harmed someone - her innocent intention exonerating her, while the harmful outcome incriminating her. Different people solve this conflict differently, suggesting the presence of dispositional moderators of the way the conflict is processed. In the present research, we explore how reasoning ability and cognitive style relate to how people choose to resolve this conflict and judge accidental harms. We conducted three studies in which we made vary the reasoning measures and the population. The results showed that individual differences in reasoning ability and cognitive style predicted severity of judgments in fictitious accidental harms scenarios, with better reasoners being less harsh in their judgments. Internal meta-analysis confirmed that this effect was robust only for accidental harms. We discuss the importance of individual differences in reasoning ability in the assessment of accidental harms.

**Keywords:** Third-party moral evaluation; Accidental transgressions; Reasoning and Thinking Style; Dual-Process toolbox

## Introduction

In 2010, the rock band *Lamb of God* was performing in Czech Republic and, during the performance, the lead singer Randy Blythe threw a fan named Daniel Nosek off the stage, with the expectation that other people will catch him. Nosek instead fell backwards directly on his head, suffered severe traumatic brain injury, slipped into a coma, and died weeks later from his injuries. If we were to be part of the jury who was going to decide how morally bad Blythe's behavior was and how much we should punish him, how would we go about it? Will we focus on his innocent intentions and reasonable beliefs about how things should have unfolded? Or will we be swayed by a strong emotional reaction in response to details about suffering that Nosek had to endure because of Blythe's actions? Will deliberating about the situation help us subdue influence of this emotional reaction on our decisions?

### *Intention and outcome in moral judgment*

When it comes to evaluating third-party harmful behavior like this, past work has shown that people rely not only on the assessment of the mental state of the perpetrator, but also on the presence of a harmful consequence for the victim[1–4]. In other words, after witnessing harmful event, a third-party moral judge reasons about the actor's intentions ("What was Blythe *thinking* when he threw his fan off the stage?!") and the victim's feelings ("How *painful* it must have been for Nosek to suffer a head trauma?"). Not only has the past work validated this two-part template of intent-based morality at the psychological level, but also explored the neural substrates for these two independent processes. In particular, this work reveals that the observers decode intentional status of interpersonal harmful actions via a network of brain regions–known as the Theory of Mind network–involved in representing others' thoughts [5], while representing victim's feeling states recruits the "empathy for pain" network[6]. Probably the most salient way to demonstrate the dissociable contributions of these two processes towards moral evaluations is by focusing on how people judge accidents. Accidental harms

elicit a strong conflict in the observer/judge because the two processes conflict with each other in terms of their output: the intent-based process focuses on innocent intentions of the actor and *reduces* severity of moral evaluations[7,8], while the outcome-based process localizes on empathic reaction towards the victim suffering and the agent's causal role in producing this outcome and *increases* severity of moral condemnation[6]. As a result, how we judge accidents depends on how we resolve the conflict posed by these two processes: difficulties in processing intentions leads to more punitive attitudes (e.g., autistic individuals[9]), while deficits in empathic reaction towards the victim can lead to forgiving attitudes (e.g., psychopathy and sadism[10,11]). In other words, forgiving accidental harms *feels* so difficult because it involves overriding a potent emotional reaction to victim suffering with a more deliberative response stemming from reasoning about intentionality. It is worth noting that this conflict is specific to accidents and is *not* encountered while evaluating other interpersonal interactions, e.g. when the actor intentionally harms someone, two processes agree on condemnation[12].

Although past work has thoroughly explored the processes that give rise to conflict while pondering over accidents and the role of dispositional mentalizing and empathizing abilities in resolving this conflict, much less attention has been paid to how one's ability and willingness to engage in analytical reasoning affect this conflict resolution. Indeed, a hint for such a role comes from work with sacrificial moral dilemmas. Sacrificial moral dilemmas, like accidents, pose a conflict of a different variety–between the emotionally aversive utilitarian option of personally harming someone and the option of letting a greater number of individuals get hurt. Past work shows that differences in individual's ability to reason and availability of cognitive resources play a role in how this conflict is resolved (for a review, see [13,14]).

***The dual-process model of thinking and its experimental toolbox***

A decade long research in the field of moral psychology has revealed how our moral judgments are, at the broadest level, the result of an interplay between emotions on the one

hand, and reason on the other. Greene and collaborators have proposed a dual-process model of moral judgment[15] that relies heavily on the field of reasoning psychology[16,17]. Although there are many variants of dual process models[16,18–21], the generic version distinguishes a fast, parallel, and almost automatic thinking system (*intuitive system*) from a slow, sequential, and cognitively effortful thinking system (*analytical system*). Note that reasoning psychologists use words interchangeably to define the two systems (e.g., intuitive/heuristic system versus analytical/deliberative/reflective system). For the sake of consistency, we will always use the words *intuitive* system/participant vs. *analytical* system/participant.

Considering that the human mind is composed of two thinking systems has led researchers to design specific manipulations and measures to test predictions derived from the dual-process model, and these protocols have proven to be useful in the field of moral judgment and decision making[13]. Specifically, numerous tasks have been designed that typically make an intuitive response conflict with an analytical response. Hence, considering the properties of the two systems, one can make a few predictions:

- people disposed (both in terms of ability and motivation) to rely on the analytical system would give a greater number of responses based on that system (*interindividual differences*);
- responses based on the intuitive system would be faster than responses based on the analytical system (*response time measure*);
- preventing people from thinking too long would decrease responses based on the analytical system (*time pressure manipulation*);
- preventing people from relying on the analytical system would decrease responses based on that system (*interfering cognitive load manipulation*).

This corpus of measures and manipulations, validated in the field of thinking and reasoning, has made it possible to characterize the way emotions (often conflated with

intuition) and reason both take part in the formation of moral judgments. This distinction has been thoroughly explored in the context of sacrificial moral dilemmas. In such dilemmas, one must choose between actively killing someone to save a greater number of persons, or not doing anything, with the consequence of letting that greater number of persons killed. Interestingly, the studies testing the predictions from the dual-process theory in the context of sacrificial moral dilemmas have used the exact same tools that are the workhorses in the reasoning literature. For example, preventing people from using their cognitive resources, using cognitive interfering tasks, delays utilitarian responses[22] and decreases the proportion of utilitarian responses[23–25]. Time pressure has been observed to selectively decrease the frequency of utilitarian responses as well [26,27], although these claims have not gone unchallenged [28]. Additionally, people who score higher on self-report measures or performance measures of reflective reasoning also tend to be more "utilitarian"[29]. Priming manipulations have also proven effective. For example, priming participants with an analytical mode of thinking made them more utilitarian than control participants, while placing participants in an experiential mode decreased their propensity for utilitarianism[30]. Finally, inducing people to think about their death (a manipulation shown to be successful in cutting into people's cognitive resource availability[31]) decreases utilitarian propensity[25]. It is worth noting, however, that there are studies which have found no effects of some of these manipulations on moral judgment activities[32], or which reanalyze previous work to highlight their limited generalizability[33,34].

Although the cross-pollination between the fields of moral and reasoning psychology has produced interesting insights about human moral psyche, most of these insights have been restricted to the domain of judgments about sacrificial moral dilemmas. Much less work has been done with other types of moral evaluations[35,36]; in particular, and of interest to us, with intent-based moral judgments. To our knowledge, only one study thus far has explored how putting people under cognitive load affects the processing of intent and consequence

information in the service of moral judgments[37]. This work revealed that under concurrent task taxing available executive resources, people relied less on harmful intent and more on harmful outcomes.

In the current work, we plan to extend the existing work to explore more broadly the role of analytical reasoning in resolving the kind of conflict one encounters while evaluating behavior of unintentional harm-doers. In a manner reminiscent of how reasoning bolsters utilitarian inclinations on moral dilemmas, we predict that more capable reasoners or people prone to analytical reasoning will experience greater ease to monitor the cognitive conflict and resolve it by overriding the strong emotional response, which would lead in turn to a greater acceptability of accidental harms. Our studies only use the interindividual differences approach.

**General Methods**

Across all studies, experimental stimuli consisted of intent-based moral vignettes[38] that were result of a 2 × 2 within-subjects design where the factors *belief* (neutral, negative) and *outcome* (neutral, negative) were independently varied such that agents in the scenario produced either a neutral outcome or a harmful outcome while acting with the belief that they were causing either a neutral outcome or a harmful outcome. The magnitude of harm severity varied freely across scenarios from mild to severe to fatal injuries. Unless otherwise stated, no time limits were imposed on responses.

We provide below an example of the parametric variation in a single scenario context in accidental harm condition:

*Background*

Matilda is walking by a neighbour's swimming pool when she sees a child about to dive in.

*Foreshadow*

The child is about to dive into the shallow end and smack his head very hard on the concrete bottom of the pool.

*Belief*

(*neutral*) Because of a label on the side of the pool, Matilda believes that the child is about to dive safely into the deep end and swim around.

*Outcome*

Matilda walks by, without saying anything to the child. The child dives in and breaks his neck.

What varied across studies was the number of scenarios used and the type of question asked, along with the scale used for measuring a response. These differences across studies are tabulated in Table 1.

***Sampling stopping rule and exclusion criteria***

These studies were each part of prior unrelated data collection, thus impeding any sampling size control in the present research.

For Amazon Mechanical Turk studies, following exclusion criteria were applied to leave out participants who: did not complete the entire survey, reported to be less than 18 years old or more than 100 years old, failed attention checks, completed the same survey multiple times. Additionally, we used TurkPrime to make the survey available for completion only to MTurk workers who had a rating of above 95% and had completed at least 100 other HITs. All sample sizes reported below refer to the *final* sample after exclusion criteria had been applied and not to the number of people who initiated their participation in the survey.

***Data availability statement***

Data and analysis scripts are available from the Open Science Framework: https://osf.io/ayb7d/

### *Data analysis*

Statistical analysis was conducted using R programming language (https://www.r-project.org/). Since the behavioral data (items within conditions within participants) had multilevel or nested structure, we utilized mixed-effects models to correctly handle the inherent dependencies in nested designs and to reduce probability of Type I error due to reduced effective sample size[39–41].

When null hypothesis significance testing (NHST) results in a failure to reject the null hypothesis (H0), this cannot be taken as evidence in support of the null hypothesis, because *p*-values are unable to quantify support in favor of the null[42]. Therefore, Bayes Factors (BF) were calculated for group comparisons to assess the relative likelihood of the null and alternative (H1) hypotheses[43]. A $BF_{01}$ of greater than 1 implies that the data are more likely to occur under H0 than under H1. Similarly, a $BF_{01}$ lower than 1 indicates that the data are more likely to occur under H1 than under H0. Thus, if we analyze data and find that $BF_{01} = 3$, this means that the data are 3 times more likely to have occurred under H0 than under H1. Based on prior guidelines[44], BFs between 1 and 3, between 3 and 10, and larger than 10 are interpreted as ambiguous, moderate, and strong support, respectively. Note that, where relevant, we provide natural logarithm values for Bayes Factors ($\log_e(BF_{01})$), which need to be exponentiated to get the $BF_{01}$.

### *Meta-analysis*

Our individual difference studies were not designed to characterize a detailed pattern of associations between reasoning measures and intent-based moral judgment (e.g., some specific correlations would be stronger than others), but instead to firmly establish the *general* form of this association. Therefore, we carried out random-effects meta-analyses[45] using regression estimates (and the associated standard errors) across measures for each study and assessed if the meta-analytic effect was significantly different than 0. In addition to providing

details from null hypothesis significance testing (NHST) approach, we also compute Bayes Factors for random-effects meta-analysis using default priors from *metaBMA* R package[46].

***Data reporting***

For the sake of brevity, results from statistical analyses are included in the figures rather than the main text (an approach adopted in the R package *ggstatsplot*[47,48]).

Similarly, demographic details for all studies (age summary statistics and gender breakdown) and details about experimental design for the studies are provided in Table 1. Additionally, more exhaustive details about the questionnaires used are provided in Supplementary Text S1, while the detailed text of the scenarios used are reported in Supplementary Text S2.

*Table 1*. Details about experimental designs included in Studies 1-3 and demographic details for participants.

| **Study** | **1** | **2** | **3** | | |
|---|---|---|---|---|---|
| **Platform used** | Lab-based | Lab-based | MTurk | | |
| **Sample size** | 44 | 112 | 1223 | 1212 | |
| **Average age** | 23.01 | 24.12 | 36.8 | 36.2 | |
| **Gender composition (% female)** | 66% | 62.50% | 50.12% | 52.56% | |
| **Country** | Italy | Italy | USA | | |
| **Measure/ Manipulation** | NFC | CRT (6-item) | REI | AOT | BB |
| **Cronbach's alpha** | 0.873 | - | 0.940 | 0.740 | 0.660 |
| **Number of conditions** | 4 | 4 | 4 | | |
| **Number of stories per participant** | 36 | 36 | 4 | | |

| | | | | |
|---|---|---|---|---|
| **Total number of stories used** | 144 (4 × 36) | 144 (4 × 36) | 16 (4 × 4) | |
| **Number of data points (after exclusion)** | 3052 | 7813 | 4892 | 4848 |
| **Type of ratings and their assignment** | Acceptability and Blame; within-subject | Acceptability and Punishment; within-subject | Wrongness and Punishment; between-subjects | |
| **Scale** | Likert | VAS | Likert | |
| **Range of ratings** | 1-7 | 0-20 | 1-7 | |

*Note*: Questions and scale labels used for different questions:
Wrongness (How wrong was [the agent]'s behavior?; 1: *not at all*, 7: *very much*), Punishment (How much should [the agent]'s be punished?; 1 (or 0): *none at all*, 7 (or 20): *a lot*), Acceptability ("How morally acceptable was [the agent]'s behavior?"; 1 (or 0): *not at all acceptable*; 7 (or 20): *completely acceptable*), Blame: "How much blame does [the agent] deserve?" (1: *none at all*, 7: *very much*).
*Abbreviations*: AOT: Actively Open-Minded Thinking, BB: Belief Bias, CRT: Cognitive Reflection Test, NFC: Need for Cognition, REI: Rational-Experiential Inventory, VAS: Visual Analog Scale

### *Ethics statement*

Across all studies, participants provided written informed consent before any study procedure was initiated. The studies conducted in Italy (1 and 2) were approved by the Ethics Committee of *Scuola Internazionale Superiore di Studi Avanzati* (Trieste) and the Hospital 'Santa Maria della Misericordia' (Udine), respectively. The study carried out on Amazon Mechanical Turk was approved by the Ethics Committee of Harvard University. All studies were conducted according to the principles in the Declaration of Helsinki.

We first investigated the relationship between cognitive ability measures and severity of intent-based moral judgments. We predicted that individuals with superior reasoning abilities would be more lenient in terms of their condemnation of unintentional harms.

**Participants.** See Table 1.

**Measures.** The following questionnaires were included in the study (for more details, see Supplementary Text S2):

- *Need for Cognition* (NFC[49]), which assesses the degree to which individuals are intrinsically motivated to engage in cognitive deliberation.
- *Cognitive Reflection Test* (6-item CRT[50]), which captures people's ability to override an appealing but incorrect intuitive response.
- *Rational Experiential Inventory* (REI[51]), which assesses the degree to which people engage in two modes of thinking: a fast, intuitive automatic thinking and a slower logical thinking.
- *Actively Open-Minded thinking* (AOT[52,53]), which assesses individual differences in disposition to consider different conclusions even if they go against one's own initial conclusion, to spend enough time on a problem before giving up, and to consider the opinions of others in forming one's own opinions
- *Belief Bias* (BB[54]), which measures the tendency to judge the strength of arguments based on the believability of their conclusion rather than how strongly they logically support that conclusion. Only syllogisms in which conclusions are logically invalid but believable (the class of problems that elicit high belief bias) are employed. They are taken from previous studies[53 55,56].

**Meta-analytic results**

Combining results across regression estimates from different studies revealed significant negative meta-analytic summary effect only for neutral and accidental harm cases such that people who scored higher on reasoning measures were more lenient in their assessment of such cases (see Figure 1). But a more careful look at the Bayes Factor for the neutral case reveals that the evidence in favor of the *alternative* hypothesis was inconclusive ($BF_{10}$ = 1.15), while it was substantial for the accidental harm cases ($BF_{10}$ = 5.15). Looking at Bayesian meta-analysis, we could also show that there was strong evidence in favor of the *null* hypothesis for the attempted ($BF_{01}$ = 17.46) and intentional ($BF_{01}$ = 24.28) harm cases, i.e., summarizing across measures, there is no relationship between reasoning ability and tendency

to judge attempted or intentional harm cases. Therefore, the meta-analysis supports our claim that scoring higher on reasoning measure is associated with greater tendency to judge third-party harmful transgressions more leniently, but only when the harm is caused accidentally.

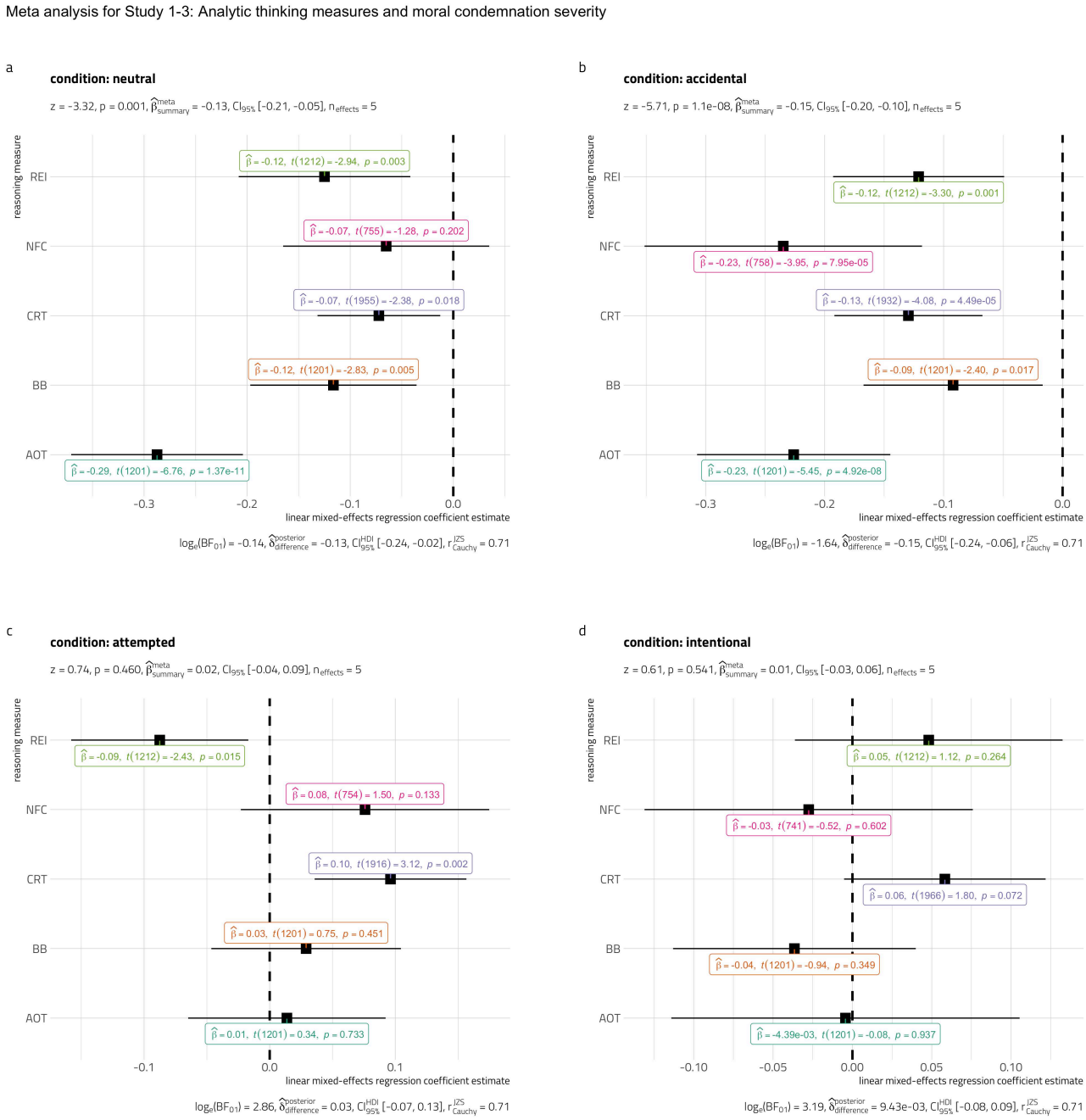



*Figure 1*. Regression coefficients for analytic thinking measures from linear mixed-effects regressions analyses carried out separately for each type of harm and each reasoning measure. The only type of harm where the regression coefficient was significantly and consistently different from 0 across all the measures was accidental

harm. Error bars indicate 95% confidence intervals. Results from frequentist random-effects meta-analysis are shown in the subtitle, while results from Bayesian random-effects meta-analysis are shown in the caption.

## General discussion

Across a series of studies, we investigated the role that reasoning in determining the severity of moral judgments about harmful transgressions. We hypothesized that reasoning would be a significant predictor of moral severity for accidental harms in particular, a class of transgressions that triggers a conflict between the outputs of two competing analyses: the mental-state based analysis focusing on the innocent intentions of the actor, which reduces severity of the moral evaluation, and the outcome-based process localizing on the victim suffering and the agent's causal role in causing the outcome, which increases severity of the moral evaluation.

Across studies, we observed that participants who self-reported to be more analytic and adept at cognitive deliberation by disposition were consistently more lenient in their judgments of accidental harms, as compared to participants who reported to rely more on the intuitive style of thinking. These results suggest that more capable reasoners (or reasoners more prone to engage in reflection) might solve the conflict posed by agent-based, intent-drive response to forgive[7,8] and victim-based, empathy-driven impulse to condemn[6] differently than less capable reasoners (or reasoners more prone to engage in intuition).

There are (at least) two possible ways in which reasoning can lead to a more lenient assessment of accidental harm-doers: (*i*) Individuals with better cognitive abilities also have more executive resources needed to be better at Theory of Mind[57], i.e. they are better at representing innocent mental states of the agent who accidentally harmed someone and thus forgive them. (*ii*) Individuals with higher propensity for cognitive deliberation are also better at down-regulating their empathic arousal stemming from harm appraisal and are thus more likely to forgive accidental harm-doers[58]. Future work should explore if it's the cognitive

(Theory of Mind) or the affective (empathic arousal) route or both that mediate the influence of reasoning ability on third-party moral evaluation.

Taken together, the present results show that individual differences in reasoning are associated with differences in the way people cope with cognitive conflict when evaluating accidental moral transgressions. The study of cognitive conflict (its detection and resolution) in the moral judgment field is an area of research still in its infancy, and we believe that the current work is a valuable addition to this growing field and hints at a number of exciting new avenue to explore.

**Acknowledgments**

We gratefully acknowledge Micaela Zucchelli for her help in creating the Qualtrics survey for Study 3. We also thank Fiery Cushman for helpful conversations and for financially supporting Study 3, and Flora Schwartz for relevant comments on the manuscript. The publication of this article was supported by an ANR grant (ANR-19-CE28-0002) awarded to B.T.

## References

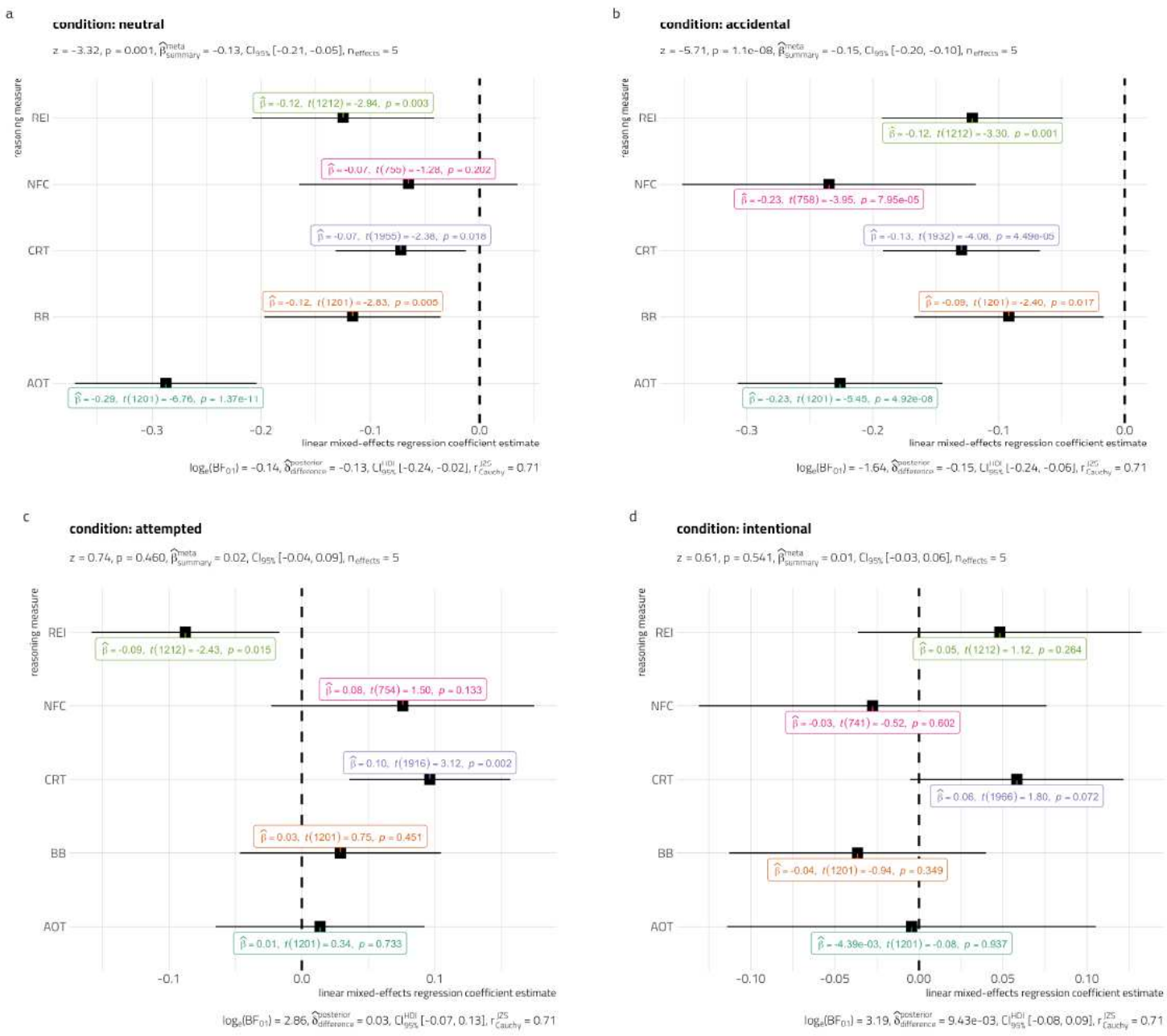
1. Cushman, F. Deconstructing intent to reconstruct morality. *Curr. Opin. Psychol.* **6**, 97–103 (2015).
2. Young, L. & Tsoi, L. When Mental States Matter, When They Don't, and What That Means for Morality. *Soc. Personal. Psychol. Compass* **7**, 585–604 (2013).
3. Buon, M., Seara-Cardoso, A. & Viding, E. Why (and how) should we study the interplay between emotional arousal, Theory of Mind, and inhibitory control to understand moral cognition? *Psychon. Bull. Rev.* **23**, 1660–1680 (2016).
4. Ginther, M. R. *et al.* Parsing the Behavioral and Brain Mechanisms of Third-Party Punishment. *J. Neurosci.* **36**, 9420 –9434 (2016).
5. Young, L., Cushman, F., Hauser, M. & Saxe, R. The neural basis of the interaction between theory of mind and moral judgment. *Proc. Natl. Acad. Sci.* **104**, 8235–40 (2007).
6. Patil, I., Calò, M., Fornasier, F., Cushman, F. & Silani, G. The behavioral and neural basis of empathic blame. *Sci. Rep.* **7**, 5200 (2017).
7. Young, L. & Saxe, R. Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia* **47**, 2065–2072 (2009).

8. Patil, I., Calò, M., Fornasier, F., Young, L. & Silani, G. Neuroanatomical correlates of forgiving unintentional harms. *Sci. Rep.* **7**, 45967 (2017).
9. Buon, M. *et al.* The role of causal and intentional judgments in moral reasoning in individuals with high functioning autism. *J. Autism Dev. Disord.* **43**, 458–70 (2013).
10. Young, L., Koenigs, M., Kruepke, M. & Newman, J. P. Psychopathy increases perceived moral permissibility of accidents. *J. Abnorm. Psychol.* **121**, 659–67 (2012).
11. Trémolière, B. & Djeriouat, H. The sadistic trait predicts minimization of intention and causal responsibility in moral judgment. *Cognition* **146**, 158–171 (2016).
12. Patil, I. & Silani, G. Alexithymia increases moral acceptability of accidental harms. *J. Cogn. Psychol.* **26**, 597–614 (2014).
13. Trémolière, B., Neys, W. De & Bonnefon, J.-F. Reasoning and moral judgment: A common experimental toolbox. in *The Routledge international handbook of thinking and reasoning* (eds. Ball, L. J. & Thompson, V. A.) 575–589 (Routledge/Taylor & Francis Group, 2018).
14. Bialek, M. & Terbeck, S. Can cognitive psychological research on reasoning enhance the discussion around moral judgments? *Cogn. Process.* **17**, 329–35 (2016).
15. Greene, J., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. The neural bases of cognitive conflict and control in moral judgment. *Neuron* **44**, 389–400 (2004).
16. Evans, J. S. B. T. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.* **59**, 255–78 (2008).
17. Kahneman, D. *Thinking Fast and Slow*. *Farrar, Straus and Giroux* (2011).
18. Epstein, S. Integration of the cognitive and the psychodynamic unconscious. *Am. Psychol.* **49**, 709–24 (1994).
19. Kahneman, D. & Frederick, S. A Model of Heuristic Judgment.ord Display. in *The Cambridge handbook of thinking and reasoning* (eds. Holyoak, K. J. & Morrison, R. G.) 167–293 (Cambridge University Press, 2005).
20. Sloman, S. A. The empirical case for two systems of reasoning. *Psychol. Bull.* **119**, 3–22 (1996).
21. Evans, J. S. B. T. & Stanovich, K. E. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* **8**, 223–241 (2013).
22. Greene, J., Morelli, S. A., Lowenberg, K., Nystrom, L. E. & Cohen, J. D. Cognitive Load Selectively Interferes with Utilitarian Moral Judgment. *Cognition* **107**, 1144–1154 (2008).
23. De Neys, W. & Białek, M. Dual processes and conflict during moral and logical reasoning: A case for utilitarian intuitions? in *Moral Inferences* (eds. Trémolière, B. & Bonnefon, J. F.) 123–136 (Psychology Press, 2017).
24. Conway, P. & Gawronski, B. Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. *J. Pers. Soc. Psychol.* **104**, 216–35 (2013).
25. Trémolière, B., Neys, W. De & Bonnefon, J.-F. Mortality salience and morality: thinking about death makes people less utilitarian. *Cognition* **124**, 379–84 (2012).
26. Cummins, D. D. & Cummins, R. C. Emotion and deliberative reasoning in moral judgment. *Front. Psychol.* **3**, 328 (2012).
27. Suter, R. S. & Hertwig, R. Time and moral judgment. *Cognition* **119**, 454–8 (2011).
28. Gürçay, B. & Baron, J. Challenges for the sequential two-system model of moral judgement. *Think. Reason.* **23**, 49–80 (2017).
29. Patil, I. *et al.* Reasoning supports utilitarian inclinations while resolving moral dilemmas. *J. Personal. Soc. Psychol.* (2020) doi:10.1037/pspp0000281.
30. Kvaran, T., Nichols, S. & Sanfey, A. The effect of analytic and experiential modes of thought on moral judgment. *Prog. Brain Res.* **202**, 187–96 (2013).

31. Trémolière, B., De Neys, W. & Bonnefon, J.-F. The grim reasoner: Analytical reasoning under mortality salience. *Think. Reason.* **20**, 333–351 (2014).
32. Tinghög, G. *et al.* Intuition and Moral Decision-Making - The Effect of Time Pressure and Cognitive Load on Moral Judgment and Altruistic Behavior. *PLoS One* **11**, e0164012 (2016).
33. Baron, J. & Gürçay, B. A meta-analysis of response-time tests of the sequential two-systems model of moral judgment. *Mem. Cognit.* **45**, 566–575 (2017).
34. McGuire, J., Langdon, R., Coltheart, M. & Mackenzie, C. A reanalysis of the personal/impersonal distinction in moral psychology research. *J. Exp. Soc. Psychol.* **45**, 577–580 (2009).
35. Callan, M. J., Sutton, R. M. & Dovale, C. When deserving translates into causing: The effect of cognitive load on immanent justice reasoning. *J. Exp. Soc. Psychol.* **46**, 1097–1100 (2010).
36. Schulz, J. F., Fischbacher, U., Thöni, C. & Utikal, V. Affect and fairness: Dictator games under cognitive load. *J. Econ. Psychol.* **41**, 77–87 (2014).
37. Buon, M., Jacob, P., Loissel, E. & Dupoux, E. A non-mentalistic cause-based heuristic in human social evaluations. *Cognition* **126**, 149–55 (2013).
38. Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A. & Saxe, R. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc. Natl. Acad. Sci.* **107**, 6753–6758 (2010).
39. Aarts, E., Verhage, M., Veenvliet, J. V, Dolan, C. V & van der Sluis, S. A solution to dependency: using multilevel analysis to accommodate nested data. *Nat. Neurosci.* **17**, 491–496 (2014).
40. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
41. Judd, C. M., Westfall, J. & Kenny, D. A. Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *J. Pers. Soc. Psychol.* **103**, 54–69 (2012).
42. Wagenmakers, E.-J. A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* **14**, 779–804 (2007).
43. Jarosz, A. F. & Wiley, J. What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *J. Probl. Solving* **7**, (2014).
44. Etz, A. & Vandekerckhove, J. A Bayesian Perspective on the Reproducibility Project: Psychology. *PLoS One* **11**, e0149794 (2016).
45. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis*. (John Wiley & Sons, Ltd, 2009). doi:10.1002/9780470743386.
46. Heck, D. W., Gronau, Q. F. & Wagenmakers, E.-J. metaBMA: Bayesian Model Averaging for Random and Fixed Effects Meta-Analysis (version 0.3.9). *CRAN* (2017).
47. Patil, I. ggstatsplot: 'ggplot2' Based Plots with Statistical Details. (2018).
48. Patil, I. statsExpressions: R Package for Tidy Dataframes and Expressions with Statistical Details. *PsyArxiv* (2021).
49. Cacioppo, J. T., Petty, R. E. & Kao, C. F. The efficient assessment of need for cognition. *J. Pers. Assess.* **48**, 306–7 (1984).
50. Finucane, M. L. & Gullion, C. M. Developing a tool for measuring the decision-making competence of older adults. *Psychol. Aging* **25**, 271–88 (2010).
51. Pacini, R. & Epstein, S. The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *J. Pers. Soc. Psychol.* **76**, 972–87 (1999).
52. Baron, J. Why Teach Thinking?-An Essay. *Appl. Psychol.* **42**, 191–214 (1993).

53. Baron, J., Scott, S., Fincher, K. & Emlen Metz, S. Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *J. Appl. Res. Mem. Cogn.* **4**, 265–284 (2015).
54. Evans, J. S. B. T., Barston, J. L. & Pollard, P. On the conflict between logic and belief in syllogistic reasoning. *Mem. Cognit.* **11**, 295–306 (1983).
55. Thompson, V. & Evans, J. S. B. T. Belief bias in informal reasoning. *Think. Reason.* **18**, 278–310 (2012).
56. Morley (née Lambell), N. J., Evans, J. S. B. T. & Handley, S. J. Belief Bias and Figural Bias in Syllogistic Reasoning. *Q. J. Exp. Psychol. Sect. A* **57**, 666–692 (2004).
57. Wade, M. *et al.* On the relation between theory of mind and executive functioning: A developmental cognitive neuroscience perspective. *Psychon. Bull. Rev.* (2018) doi:10.3758/s13423-018-1459-0.
58. Treadway, M. T. *et al.* Corticolimbic gating of emotion-driven punishment. *Nat. Neurosci.* **17**, 1270–5 (2014).

# Figures




## Figure 1

Regression coefficients for analytic thinking measures from linear mixedeffects regressions analyses carried out separately for each type of harm and each reasoning measure. The only type of harm where the regression coefficient was significantly and consistently different from 0 across all the measures was accidental harm. Error bars indicate 95% confidence intervals. Results from frequentist randomeffects meta-analysis are shown in the subtitle, while results from Bayesian randomeffects meta-analysis are shown in the caption.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- reasoningintentbasedjudgmentssupplementary.pdf