



# How does reasoning influence intentionality attribution in the case of side effects?

Nicola Matteucci Armandi Avogli Trotti<sup>1</sup> · Micaela Maria Zucchelli<sup>1</sup> · Andrea Pavan<sup>1</sup> · Laura Piccardi<sup>2,3</sup> · Raffaella Nori<sup>1</sup>

Received: 4 April 2024 / Accepted: 8 August 2025 / Published online: 21 August 2025  
© The Author(s) 2025

## Abstract

To evaluate others' actions objectively, one must integrate the actor's mental states with the potential consequences of his actions. However, consequences can distort the perception of intentionality. The Knobe effect, or “side-effect effect,” demonstrates that individuals attribute greater intentionality to negative than positive foreseen yet unintended side effects. This study explores how reasoning styles and abilities influence these judgments. A sample of 172 college students completed validated reasoning style questionnaires, including the Rational-Experiential Inventory (REI) and the Actively Open-Minded Thinking scale (AOT), a syllogistic reasoning task, and scenario-based tasks in a randomized, between-subjects design (negative vs. positive side effect). Our findings reveal that a more deliberative reasoning style and longer response times both reduce bias in attributing intentionality to negative side effects, highlighting two distinct pathways through which response times mediate the influence of reasoning style on reducing biased judgments. We explore how reasoning affects our attributions of intentionality leading to a more balanced consideration of an actor's mental state and the consequences in moral judgment.

## Introduction

Living in complex societies, human beings have evolved various general cognitive processes to navigate the challenges of intricate social interactions. Humans use these processes when interacting with each other and observing situations from an outsider's perspective (Buckholtz et al. 2012). Most legal systems rely on judges' impartial judgment and reasoned decisions (Darley 2009; Robinson et al. 2007). These decisions depend on two important factors: the

mental state of the person while acting (*mens rea*) and the consequences of the action (*actus reus*) (Cushman 2008), which influence verdicts and sentences (Shen et al. 2011). Sometimes, *mens rea* and *actus reus* do not align, creating a cognitive challenge. This is especially true in cases of accidental harm, recklessness, or negligence, where evaluators must judge individuals who caused negative outcomes without intent to harm (Margoni and Surian 2022; Patil and Trémoлиère 2021; Young and Saxe 2009).

The Knobe effect in literature describes how intent (*mens rea*) and action (*actus reus*) influence the perception of intentionality based on the type of outcome (Knobe 2003). The Knobe effect shows that people consider foreseeable but unintended negative side effects of an action to be more intentional than positive ones. This highlights how consequences influence judgments of intentionality, shaping whether an action is seen as more intentional or accidental. The effect is robust across cultures, age groups, and scenarios (Burra and Knobe 2006; Feltz 2007; Leslie et al. 2006). Many studies have tried to understand and explain the imbalance in assigning intentionality between negative and positive side effects, exploring many aspects, such as moral valence, emotional reactions to negative results, and thinking about the actor's thoughts (Feltz 2007; Knobe

---

Editors: Riccardo Brunetti (European University of Rome), Allegra Indraco (European University of Rome); Reviewers: two researchers who prefer to remain anonymous

✉ Nicola Matteucci Armandi Avogli Trotti  
nicola.matteucci4@unibo.it

✉ Raffaella Nori  
raffaella.nori@unibo.it

<sup>1</sup> Department of Psychology, University of Bologna, V.le Berti Pichat, 5, Bologna, Bo 40127, Italy

<sup>2</sup> Department of Psychology, “Sapienza” University of Rome, Rome, RM, Italy

<sup>3</sup> IRCCS San Raffaele Cassino, Cassino (FR), Italy

2022; Zucchelli et al. 2018, 2019). Moral valence is one of the most widely accepted theories (Knobe 2003). This theory suggests that moral judgments affect how individuals perceive intentionality. When a side effect is perceived as morally negative, people are more likely to think it was intentional because they believe the person deserves to be blamed for its occurrence.

Previous research suggests that evaluating accidental harm takes longer than assessing intentional harm, as individuals need extra time to evaluate and integrate conflicting information about mental state (*mens rea*) and consequences (*actus reus*) into their judgments (Buckholtz et al. 2008; Buon et al. 2016; Decety and Cacioppo 2012; Young and Saxe 2009). Dual-Process theories (DPT) offer a framework for understanding these judgments, distinguishing between an intuitive system (System 1) that is fast and emotion-driven, and a deliberative system (System 2) that is slower and based on analytical reasoning (Evans 2008; Kahneman 2011; Sloman 1996). The two-process model of intent-based morality posits that judgments of accidental harm can be seen as the outcome of a cognitive conflict between an outcome-based process (System 1), influenced by emotional reactions to harm severity, and an intent-based process (System 2) which seeks to exculpate the agent by integrating information about their benign mental state (Cushman 2008, 2013; Cushman et al. 2006). Research shows that slower, controlled reasoning can resolve the cognitive conflict in evaluating accidental harm, leading to more lenient moral judgments. Disrupting this process with time constraints or cognitive load results in harsher judgments (Buon et al. 2013; Martin et al. 2021; Schwartz et al. 2024), while individuals with higher rational reasoning style evaluate accidental harm more leniently (Patil and Trémolière 2021; Schwartz et al. 2022). The authors hypothesized that these results arise because better reasoners are more able to prioritize *mens rea* over the more intuitive and emotional *actus reus* information when making judgments. Similarly, also reasoning abilities have been found to be a good predictor of individuals' capacity to override intuitive and spontaneous responses in accidental harm evaluation (Patil and Trémolière 2021) or moral dilemmas (Patil et al. 2021) following the same pattern described for the rational reasoning style. Ngo et al. (2015) explored the neural basis of the Knobe effect, finding that the amygdala's higher activity correlates with stronger attributions of intentionality in negative situations, driven by subjective emotional reactions. They suggested that the dlPFC is involved because it plays a role in the cognitive theory of mind, but found no mediating effect on subjective emotional responses. Their study found that participants reacted more quickly to negative scenarios, suggesting that stronger emotional responses drive the Knobe effect. In support of this, Pinillos et al.

(2011) found that individuals with higher cognitive reflection abilities exhibited a weaker Knobe effect, suggesting how deliberative reasoning can mitigate its influence.

Building on the existing findings from the literature on moral evaluation of accidental harm, we aim to investigate the potential role of reasoning in diminishing the asymmetry in intentionality attribution seen in the Knobe effect. To do this, we will evaluate participants' reasoning styles and abilities, as these factors may help alleviate the biases associated with *mens rea* and *actus reus* integration.

Rational reasoning style has been described with different facets and underlying constructs (Baron et al. 2015; Decety and Cacioppo 2012; Thompson & Evans 2012). In the moral field, the Rational-Experiential Inventory (REI; Pacini and Epstein 1999) and the Actively Open Minded Thinking scale (AOT; Stanovich and Toplak 2023; Stanovich and West 1997) are the two most commonly used measures (Trémolière et al. 2017). The REI is based on Epstein's Cognitive-Experiential Self-Theory (CEST) (Epstein et al. 1996; Pacini and Epstein 1999), grounded in dual-process theories, the REI measures individual preferences for these thinking styles through two subscales that distinguish between two cognitive systems: a fast, intuitive experiential system (measured by REI-E subscale) and a slower, deliberate and rational one (measured by REI-R subscale). The AOT scale (Baron 1993, 2005; Stanovich and Toplak 2023), measures openness to alternative viewpoints, sensitivity to contradictory evidence, and resistance to cognitive closure. It reflects a reflective, analytical, and unbiased reasoning disposition. The construct aims for intellectual flexibility, such as considering opposing evidence, or different points of view, while discouraging dogmatism. While REI-R and AOT both center on rational thought, they differ in scope and emphasis. REI-R focuses on a cognitive disposition toward logic, problem-solving, and detailed analytical examination, whereas AOT pursues a specific style of cognitive regulation aimed at not systematically reinforcing previously acquired beliefs (Baron 2020; Stanovich and Toplak 2023; Trémolière et al. 2017).

Building on previous theories about deliberative processing in moral cognition, we propose that higher-order thinking reduces biases from emotional responses and allows for better evaluations focused on the actor's mental state. We hypothesize that individuals with stronger deliberative reasoning skills, measured by the REI, AOT, and syllogistic reasoning tasks, will be more effective at integrating *mens rea* and *actus reus* when evaluating scenarios related to the Knobe effect. These individuals should rely less on emotional reactions and more on reflective, analytical thinking, resulting in a decreased tendency to assume intent in negative side effects. This proposed mechanism aligns with previous studies showing that deliberate reasoning can counter

biased responses in moral judgments based on intent (Patil and Trémolière 2021; Schwartz et al. 2022). Moreover, we predict that response times will further reveal this dynamic: individuals who spend more time reasoning through the Knobe scenarios and engaging with the information in depth will attribute lower levels of intentionality to negative side effects compared to those who respond quickly. This aligns with the idea that slower, deliberative processes reduce reliance on automatic, emotionally driven judgments.

In summary, our key prediction is that individuals who engage in more deliberative thinking are likely to display reduced bias in scenarios related to the Knobe effect, particularly in terms of attributing intentionality to negative side effects. This leads to two main conclusions: (i) More deliberative reasoning reduces reliance on emotions and intuition; (ii) Emphasizing analytical thinking promotes a balanced view of both intent and outcomes.

## Materials and methods

### Participants

Our sample comprised 172 participants (GPower 3.1. Faul et al. 2007, 2009;  $f2=0.25$ ,  $\alpha=0.05$ , Power=0.90), recruited at University and from citizen associations through notices on social networks and bulletin boards (age  $M=21.75$ ,  $SD=2.14$ ; 31% males; education  $M=14.89$ ,  $SD=2.57$ ). Exclusion criteria included a history of neurological or psychiatric disorders, as well as the use and abuse of substances. No participants were excluded. All participants signed a written consent form before the study commenced. The local Ethics Committee approved the study (Prot. n. 130861, University of Bologna, Italy).

### Task and materials

All materials and tasks related to the experiment are detailed in the supplementary section of this paper. Additionally, the data presented in this study can be accessed through the OSF repository at the following link:

[https://osf.io/cvj5y/?view\\_only=b41ffd2cf3a542bb8fc6091c7fafa185](https://osf.io/cvj5y/?view_only=b41ffd2cf3a542bb8fc6091c7fafa185).

**Knobe effect scenario (Ngo et al. 2015; Zucchelli et al. 2018, 2019)**

We selected eight negative and eight paired positive scenarios from prior research on the Knobe effect using the same set of stimuli adopted by Zucchelli et al. (2018, 2019), and drawing from the eighty scenarios identified by Ngo et al. (2015). This selection was made because these stimuli

have proven effective and sensitive in eliciting differences in intentionality attribution in previous studies employing similar designs. Moreover, it promotes methodological consistency and comparability with earlier research. Each scenario depicts a character's action that leads to an unintended but anticipated consequence (see supplementary material for the complete set). Following each scenario, participants rated the intentionality of the side effect on a Likert scale from 0 (not intentional at all) to 7 (completely intentional).

**Rational experiential inventory (REI; Pacini and Epstein 1999)**

The REI was used to assess participants' reliance on two main modes of thinking: fast, intuitive, automatic thinking and slower, logical thinking. This experiment analyzed 20 items from the REI-R subscale related to rational thinking, such as "*I have no problem thinking things through carefully*". Participants rated how well each item described them on a 5-point scale (1=definitely not true of myself, 5=definitely true of myself). Higher scores indicated a greater degree of rationality. Several studies have provided evidence of good test reliability: *Cronbach's alpha*=0.86 to 0.90 (Björklund and Bäckström 2008; Pacini and Epstein 1999; Sladek et al. 2010; Wittman et al. 2009).

**The actively Open-Minded thinking scale (AOT; Baron et al. 2015; Baron 2019)**

This scale evaluates participants' tendencies toward open-minded thinking. As described in the introduction, this includes their willingness to fairly consider diverse conclusions, even those contradicting their initial preferences; their careful consideration of others' viewpoints when forming their own opinions; and their general propensity to actively and openly contemplate available information. Items on the AOT scale capture attitudes such as "*People should always take into consideration evidence that goes against their beliefs*" and "*Allowing oneself to be convinced by a solid opposing argument is a sign of good character*". Participants respond on a 5-point scale, from 1 = "*Completely disagree*" to 5 = "*Completely agree*", with respect to each of the eight items. Higher scores indicate a greater tendency toward an open-minded thinking style. The scale demonstrated good reliability, with *Cronbach's alpha*=0.65 to 0.92 (Baron 2019; Janssen et al. 2020; Stanovich and Toplak 2023).

**Belief bias (BB; Evans et al. 1983)**

This task measures reasoning abilities and shows the conflict between intuitive processes, which favor belief, and deliberative processes, which prioritize logical structure. In

a typical syllogistic task, participants evaluate arguments where the logical validity of the conclusion does not align with its intuitive plausibility, thus assessing logical reasoning and one's capacity to resist cognitive biases. This makes it a valuable tool for understanding the interaction between rational and intuitive cognitive processes. In a conflict syllogism, conclusions are logically invalid but believable, as in the following example:

*All quadrupeds are mammals.*

*All quadrupeds are animals.*

*Therefore, all mammals are animals.*

The ability to solve conflict syllogisms is closely related to higher reasoning abilities and engagement in reflective thinking (Baron et al. 2015; Evans et al. 1983; Klauer et al. 2000; Thompson & Evans 2012). We used eight conflict syllogisms taken from previous studies on belief bias in syllogistic reasoning (Morley et al. 2004; Thompson & Evans 2012). Participants judged whether the conclusion logically followed using a binary YES/NO response. Participants received 1 point for each correctly solved syllogism, with higher scores indicating lower belief bias (i.e., greater reliance on analytical reasoning).

## Procedure

The experiment was administered online through the software Qualtrics (First release: 2005, Provo, Utah, USA, available at: <https://www.qualtrics.com>). After providing demographic information, participants completed the experiment in a randomized order. Specifically, half of the participants first answered the Knobe Effect questions, and then completed the reasoning questionnaires (REI, AOT, and syllogisms). The other half did the reasoning questionnaires first, followed by the Knobe scenarios. The order of items in both the scenarios and reasoning questionnaires was randomized to prevent order effects. Participants were randomly assigned to one of two main conditions: negative side effects or positive side effects (*between-subjects design*; Zucchelli et al. 2019). Each scenario appeared across two screens: the first described the context and actions, while the second asked participants to rate the side effect's intentionality on a 7-point Likert scale (0 = *not intentional at all*, 7 = *completely intentional*). This design accurately measured response times for the intentionality question before confirming answers. The procedure was the same for all of the 8 scenarios for each participant. The experiment lasted approximately 30 min.

## Statistical analysis

For each participant, we calculated the mean of intentionality attribution scores and response times. Initially,

we tested for differences in intentionality attribution and response times between negative and positive scenarios using independent-samples *t*-tests. This was done to assess the presence of the Knobe effect in the selected scenarios and to compare reasoning times between the two types of scenarios. Next, we performed two mediation models - one for negative and one for positive scenarios - with reasoning measures as predictors and response times as mediators. We aimed to determine how reasoning influences intentionality attribution differently in negative and positive scenarios. Our analysis aimed to determine whether different facets of reflective thinking affect the integration of *mens rea* and *actus reus* over time. The analyses were conducted using JASP (version 0.17.2.1).

## Results

There were no significant differences in demographic variables or reasoning skills and abilities between the two groups (see supplementary materials,  $ps > 0.05$ ). We tested the normality of the residuals for the intentionality attribution scores using the Shapiro-Wilk test. The results were significant (Shapiro-Wilk = 0.94,  $p < .001$ ), indicating some deviation from normality. The skewness and kurtosis were  $-0.67$  ( $SE = 0.185$ ) and  $-0.37$  ( $SE = 0.37$ ), respectively. Given the deviation from normality suggested by the Shapiro-Wilk test, we used the Mann-Whitney *U* test to examine the differences between the two conditions. We found that negative side effects ( $M = 5.23$ ,  $SD = 0.98$ ) were perceived as significantly more intentional than positive ones ( $M = 2.82$ ,  $SD = 1.56$ ) (Mann-Whitney  $U = 628$ ,  $p < .001$ , *Rank-Biserial Correlation* = 0.83). Besides, response times were significantly faster for negative side effects ( $M = 3.57$  s,  $SD = 3.44$  s) compared to positive ones ( $M = 4.32$  s,  $SD = 2.23$  s) (Mann-Whitney  $U = 2380$ ,  $p < .001$ , *Rank-Biserial Correlation* = 0.36).

## Negative scenarios mediation model

All the other measures significantly correlated with each other, except for syllogisms (see Table 1), which were not found to predict intentionality attribution and were therefore excluded from the model.

We consequently performed a mediation model with REI and AOT scores as predictors, the mean of response times as mediating factor, and the mean of intentionality attribution as the dependent variable. We used bootstrapping (1,000 samples with 95% bias-corrected confidence intervals) as a non-parametric resampling procedure to estimate the significance of the indirect effect. A significance level of 0.05 was

**Table 1** Pearson's correlation coefficients

	Intentionality scores	Reaction times	AOT	REI	Syllogisms
Intentionality scores	-				
Reaction times	-0.411***	-			
AOT	-0.319**	0.241*	-		
REI	-0.335**	0.330*	0.256*	-	
Syllogism	-0.200	0.149	0.089	0.061	-

\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$

used to determine the statistical significance of both direct and indirect paths.

The total effect of REI and AOT were both significant ( $\beta = -0.016$ ,  $p < .01$ , 95% CI  $[-0.031, -0.001]$ ) and ( $\beta = -0.064$ ,  $p < .05$ , 95% CI  $[-0.115, 8.2 \times 10^{-4}]$ ), respectively. The mediation analysis showed that response times significantly and fully mediated the effect of REI scores on intentionality judgments for negative side effects ( $\beta = -0.005$ ,  $p < .05$ , 95% CI  $[-0.015, -5.439 \times 10^{-4}]$ ). However, this was not the case for the AOT scores, where no significant mediation was observed ( $\beta = -0.013$ ;  $p = .156$ , 95% CI  $[-0.049, 0.003]$ ). The mediation for the REI scores was complete, as the direct effect of REI on intentionality judgments become not significant when response times were included in the model ( $\beta = -0.011$ ;  $p > .05$ , 95% CI  $[-0.024, 0.002]$ ). Additionally, we observed a negative relationship in the direct path from intentionality attribution to response times ( $\beta = -0.302$ ;  $p < .01$ , 95% CI  $[-0.632, -0.072]$ ) (Fig. 1).

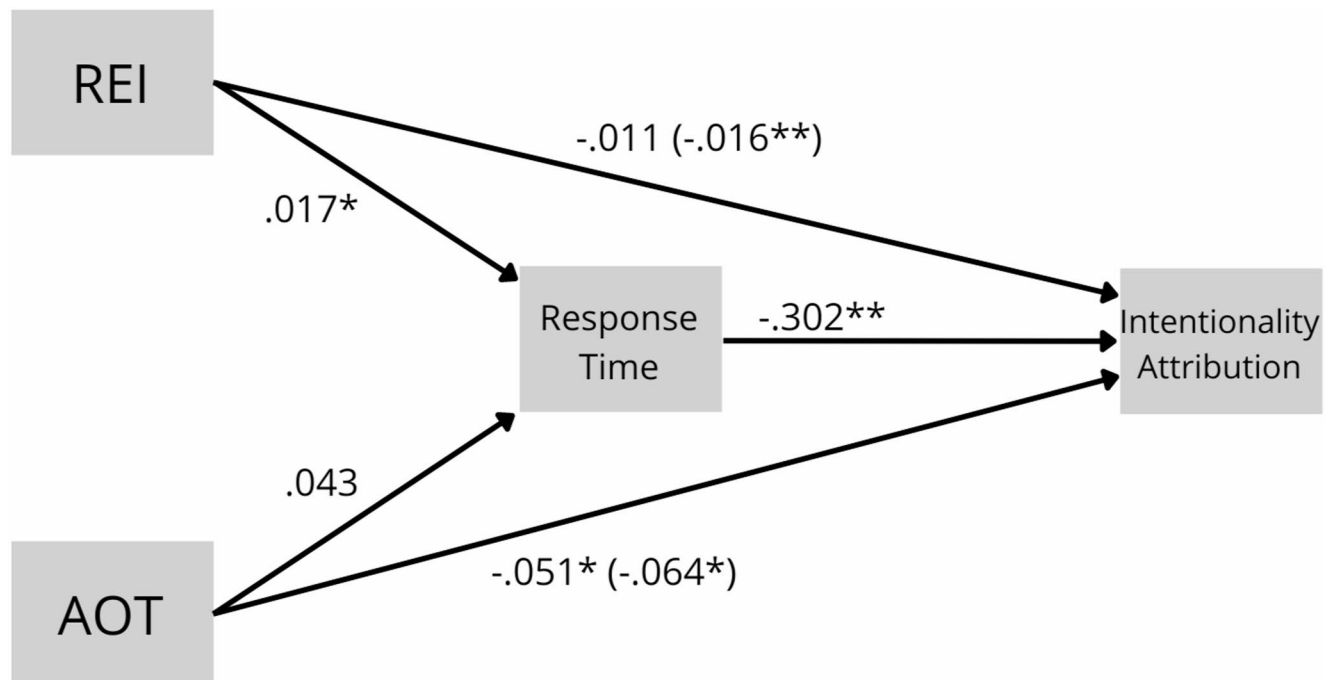
## Positive scenarios mediation model

No significant relationships were found between reasoning abilities, positive side effects, intentionality judgments, and response times.

## Discussion

The present study aimed to examine how reasoning style and abilities modulate the influence of *actus reus* (the consequences of an action) on *mens rea* (mental states/intentions) when attributing intentionality to foreseeable but unwilling positive or negative side effects (i.e. the Knobe effect). We hypothesized that people with stronger reasoning skills would attribute less intentionality to negative side effects, suggesting that those who engage in deeper reflection are more likely to focus on the absence of intention in the action itself, rather than on the negative outcomes.

Our results confirmed the presence of the Knobe effect, with higher intentionality attribution toward negative side effects (Knobe 2003). Consistent with previous findings, we also observed shorter response times when participants evaluated negative compared to positive side effects (Ngo et al. 2015). This faster response to negative side effects may reflect an emotional response triggered by the negative consequences (Nadelhoffer 2004; Ngo et al. 2015; Zucchelli et al. 2019). Relying on emotions, participants likely responded more intuitively, predominantly engaging



**Fig. 1** Mediation model showing the full mediation of response times between REI scores and intentionality attribution. Values are bootstrapped path coefficients, with total effects shown in brackets. \*  $p < .05$ , \*\*  $p < .01$

System 1 processes. This would result in a simplified mental model that relies more on consequences than on intentions, as postulated by the two-process model of moral judgment (Cushman 2008). Similarly, when participants are forced to rely on System 1 - such as under time pressure or cognitive load (Buon et al. 2013; Martin et al. 2021; Schwartz et al. 2024) - they tend to place more emphasis on consequences rather than intentions when evaluating accidental harm scenarios.

Regarding our main hypotheses, the results confirmed that both an open-minded reasoning style (AOT) and a rational reasoning style (REI-R) were associated with lower intentionality attribution to negative side effects. Besides, longer response times, indicating more extended reasoning to give their response, showed the same pattern. These three measures also correlate with each other. A higher propensity for rational thinking has been linked to less harsh judgements of accidental harm (Patil and Trémolière 2021; Schwartz et al. 2022), consistent with our findings that this propensity facilitates a more accurate consideration of *mens rea*. Thus, our results, along with previous evidence, suggest that a rational reasoning style can influence how people integrate intent-based and outcome-based judgments in moral decision-making (Cushman et al. 2006; Schwartz et al. 2022).

Interestingly, our study demonstrates that reasoning may directly influence the perception of intentions in case of negligence (i.e., foreseen but unintended consequences), whereas previous research focused primarily on moral judgments such as punishment, moral acceptability, or culpability, and has only considered cases of pure accidental harm (Buon et al. 2013; Martin et al. 2021; Patil and Trémolière 2021; Schwartz et al. 2022). Our results suggest that reasoning may influence how we evaluate events by shaping our perception of intentions even before affecting later judgments.

Although our study did not directly assess perceptions of negligence, the scenarios employed, in which agents foresee negative side effects yet choose to act regardless, represent clear examples of negligent behavior, where foreseeable consequences are knowingly disregarded. Indeed, definitions of negligence (e.g. Nunez et al. 2014; Flick et al. 2024), according to the Malle and Knobe's (1997) model of intentionality, consistently incorporate features of 'knowledge' and 'awareness', meaning that the actor knew or was aware of the potential consequences of their actions, but do not include the 'desire' to bring about negative outcomes, just like in the Knobe's negative scenarios (Guglielmo and Malle 2010).

Our results are the first to show that reasoning style can directly influence perceived intentionality in such contexts, rather than merely affecting subsequent moral evaluations.

Previous literature has, in fact, primarily focused on the role of reasoning in generating more lenient judgments in cases of purely accidental harm, without explicitly manipulating or measuring negligence (Patil and Trémolière 2021; Schwartz et al. 2022).

Nonetheless, our findings may offer even greater interpretive value in light of recent studies highlighting the spontaneous inferences of recklessness and negligence that often arise when evaluating classically accidental scenarios.

Research shows that people often attribute negligence when only intentions and outcomes are known and no explicit information is provided regarding foreseeability or care (Nobes and Martin 2022). Moreover, even describing someone as acting carefully yet still causing harm can lead individuals to infer that the agent should have foreseen the harmful consequences of their actions, thereby inferring negligence in the absence of specific information (Margoni and Surian 2022). These findings are consistent with developmental and lifespan research showing that both young children and older adults tend to over-attribute negligence in response to accidental negative outcomes, particularly when information about care or foreseeability is missing or under-specified (Margoni et al. 2019; Nobes et al. 2009, 2017). It has been suggested that properly processing and integrating information about *mens rea* requires substantial cognitive effort, which may be more challenging for these population groups due to underdeveloped or declining executive function skills (Margoni et al. 2018, 2019; Margoni and Surian 2016).

In light of our findings on reasoning propensity and intentionality attribution, we speculate that better reasoners may judge accidental harm more leniently not only because they are less influenced by outcomes in their moral evaluations, but also because, even when considering different possible levels of care and foreseeability, they ultimately perceive the action as less intentional, and thus intrinsically less blameworthy or less deserving of punishment.

Our study, which directly measured intentionality attribution rather than moral or punishment judgments in response to negative and foreseeable outcomes, combined with individual differences in reasoning, offers a novel perspective on the role of reasoning in intent-based moral judgment. Given that previous research has shown how negligence is often inferred in evaluations of accidental harm, the mechanism identified in our findings may also apply to such cases.

Specifically, better reasoners may be more likely to perceive actions involving possible negligence and a negative outcome as less intentional from the outset, rather than simply focusing less on the outcome during the construction of punishment and moral judgments.

Our results thus contribute, through the combination of reasoning propensity measures and intentionality attribution,

to adding a new piece to the puzzle of intent-based moral judgment and reasoning. They suggest that deliberative thinking can act directly on a core antecedent of moral evaluation, namely perceived intentionality (Cushman 2015), reflecting the idea, widely discussed in the Knobe effect literature, that judgments of intentionality and moral evaluation may not follow a fixed or strictly sequential order, but can interact and shape each other in dynamic ways (Feltz 2007; Leslie et al. 2006; Pettit and Knobe 2009).

Our mediation analysis, which considered REI-R and AOT as predictors and response times as a mediator, revealed important distinctions between these two aspects of rational thinking. For REI-R, longer response times fully mediated the attribution of intentionality, indicating that reflective, deliberate reasoning plays a key role in diminishing the Knobe effect. In contrast, for AOT, response times did not significantly mediate effect, suggesting that open-minded thinking operates independently of prolonged deliberation when modulating biased responses in the Knobe effect. These findings highlight distinct cognitive mechanisms through which individuals with high REI-R and AOT scores reduce the Knobe effect. Although correlated, REI-R and AOT represent different facets of rational reasoning, operating through different mechanisms. Prior studies confirm this distinction, showing that despite their correlation, they influence cognition and behavior in distinct ways (MacLaren et al. 2012; Ozono & Sakakibara 2024; Svedholm-Häkkinen and Lindeman 2017; Szasz et al. 2017).

Individuals with high REI-R scores rely on reflective, analytical reasoning, aligning with System 2 processes (Kahneman 2011; Stanovich and West 2000). This resource-intensive approach involves deconstructing information, weighing evidence, and suppressing initial intuitive responses - especially in emotionally charged contexts like the Knobe effect. This deliberate method of conflict resolution could clarify why participants with high REI-R took longer to diminish their biases. Their reasoning style, which prioritizes constructing rational, well-founded judgments, led them to focus more on the agent's lack of intent, rather than being mainly influenced by the negative consequences.

In contrast, individuals with high AOT scores adopt a flexible, integrative reasoning style, characterized by openness to different perspectives and a readiness to revise conclusions based on new information (Stanovich and Toplak 2023; Svedholm-Häkkinen and Lindeman 2017). They are more adept at processing conflicting information early, distinguishing foreseen but unintended consequences from fully intended actions. This ability likely explains why their judgments were less dependent on response time. The key difference between these reasoning styles lies in their approach; high REI-R involves effortful conflict resolution, whereas high AOT reflects a more dynamic integration.

While high REI-R individuals process information in a way that resembles the traditional dual-process framework, where System 2 overrides System 1 (Evans 2003; Evans & Stanovich 2013), high AOT individuals reflect a more integrative model, where intuitive and logical reasoning coexist dynamically, enabling them to form balanced judgments more quickly. This aligns with recent dual-process theories, which suggest that logical responses can occur intuitively without the need for explicit intervention of System 2 intervention (Bago and De Neys 2019a; Bago and Neys 2019b; De Neys and Pennycook 2019). Both reasoning styles, however, contribute to reducing the Knobe effect bias, albeit through distinct cognitive pathways.

Our findings offer insights into the coexistence of two hypothesized mechanisms within the brain's central executive network that regulate *mens rea* and *actus reus* information (Krueger and Hoffman 2016; Yang et al. 2024). Specifically, the Integration and Selection Hypothesis (Buckholz et al. 2008, 2012), suggests that the dlPFC integrates and selects relevant information during judgment, aligning with the strategy used by high AOT individuals. Conversely, the Cognitive Control Theory (Knoch et al. 2006; Spitzer et al. 2007) proposes that the dlPFC exerts top-down regulation to suppress emotional responses, consistent with the deliberate, step-by-step reasoning observed in high REI-R individuals. We did not find any significant relationship between syllogistic reasoning abilities and intentionality attribution for negative side effects. This suggests that REI-R and AOT may better capture the reasoning required to resolve cognitive conflicts between intentions and consequences in moral judgments. Unlike syllogistic reasoning, which focuses on logical validity without emotional content, intentionality attribution in the Knobe effect involves integrating emotionally charged, context-specific information about *mens rea* and *actus reus*, while managing intuitive reactions. As a result, syllogistic reasoning may be less effective in predicting the ability to inhibit intuitive and emotional responses in the Knobe effect and properly integrate morally relevant information. The absence of a moral or social dimension in syllogistic reasoning may limit its applicability to real-world scenarios that require complex emotional and theory of mind-based reasoning.

Interestingly, none of our reasoning measures were significantly related to judgments of positive side effects. This may be because positive consequences without specific intent involve a simpler evaluative process, requiring less effort to inhibit intuitions or resolve conflicting information. Consequently, higher reasoning abilities or stronger regulation of intuitive responses do not seem to influence intentionality attribution in these cases. The contrast between negative and positive side effects underscores the distinct cognitive and emotional processes involved. As previously

suggested (Ngo et al. 2015; Pinillos et al. 2011), negative side effects elicit stronger intuitive, emotionally driven responses, mediated by System 1. Individuals with a greater propensity for rational thinking are better equipped to regulate these intuitions, which explains their ability to modulate intentionality attribution in negative scenarios.

Our findings are the first to demonstrate that two distinct forms of rational thinking can lead to similar outcomes in moral judgment through different cognitive pathways. This distinction has important implications for experimental paradigms that rely on manipulations such as time pressure or cognitive load to influence reasoning processes (Buon et al. 2013; Martin et al. 2021; Zucchelli et al. 2025). Specifically, it suggests that individuals with different reasoning profiles may be differentially susceptible to such manipulations. Furthermore, our results help deepen our understanding of dual-process theories on intent-based moral judgment, showing how participants with a preference for slow and open reasoning are not only less influenced by negative outcomes in moral judgment (Patil and Trémolière 2021; Schwartz et al. 2022), but also less influenced by them when it comes to perceiving the intentions behind the action. As stated above, this opens new ways of interpreting the role of reasoning in intent-based moral judgment, and raises new interesting research questions regarding the impact of reasoning on the perception of intention, consequences, and consequent moral judgments.

### Limitations and future research

While our findings provide valuable insights into the cognitive mechanisms underlying intentionality attribution, some limitations should be noted. First, our reliance on REI-R and AOT may not fully capture the interplay between reasoning, emotional regulation, and intentionality judgments. Specifically, these measures assess general reasoning dispositions but do not directly evaluate cognitive processes such as inhibitory control, which could play a crucial role in intentionality decision-making. Future studies could incorporate additional specific measures, such as classic and emotional Stroop tasks (Greco 1993; Stroop 1935) for measuring inhibition abilities or controls for subjective emotional responses. Second, we did not manipulate reasoning directly (e.g., through cognitive load or time pressure), which limits our ability to establish a causal link between reasoning and intentionality attribution. Incorporating such manipulations could refine our understanding of the temporal dynamics of reasoning in this process. Methods like the two-response paradigms (Bago and De Neys 2019a; De Neys and Pennycook 2019; Trémolière et al. 2017), could help address this gap. Third, the null results for belief bias syllogisms suggest the need for measures that better reflect

the emotional and cognitive demands of the task. In fact, we manipulated only the type of consequences in our scenarios. Considering emotional factors and different levels of conflicting information in problem-solving can help identify when reasoning skills are most influential. Future research should explore how reasoning influences judgments in cases of accidental harm, by controlling for perceived intentionality and foreseeability. This would help differentiate between truly accidental harm and harm caused by recklessness or negligence. Expanding these methodological approaches will provide a more comprehensive understanding of the interaction between reasoning and intuitive processes in intent-based moral judgment.

### Conclusions

In summary, the present findings highlight the intricate nature of integrating *mens rea* and *actus reus* information in human cognition. In instances where these components interact, as evidenced by the Knobe effect, it becomes imperative to engage higher-order cognitive processes to regulate strong, automatic intuitive and emotionally driven reactions, or to selectively prioritize and integrate all pertinent information from the outset. The observed differences in reasoning strategies demonstrate how distinct cognitive pathways - such as deliberate, reflective suppression and flexible, integrative reasoning - can lead to similar outcomes, namely, reducing the attribution of intentionality to negative side effects.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10339-025-01300-w>.

**Author contributions** Nicola Matteucci: Writing - original draft; Methodology; Data curation; Formal analysis. Micaela Maria Zucchelli: Conceptualization, Methodology, review & editing. Andrea Pavan: Data curation; Formal analysis; review & editing. Laura Piccardi: Supervision; review & editing. Raffaella Nori: Conceptualization; Supervision; Project administration; review & editing.

**Funding** Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

### Declarations

**Conflict of interest** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Ethical approval** All procedures performed were in accordance with the ethical standards of the institutional and/or national research committee and with the 2025 World Medical Association Declaration of Helsinki. The study has been approved by the Ethic Committee of the University of Bologna (Prot. no. 130861).

**Informed consent** Each participant provided written informed consent for participation and publication.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bago B, De Neys W (2019a) The intuitive greater good: testing the corrective dual process model of moral cognition. *J Exp Psychol Gen* 148(10):1782–1801. <https://doi.org/10.1037/xge0000533>
- Bago B, De Neys W (2019b) The smart system 1: evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Think Reason* 25(3):257–299. <https://doi.org/10.1080/13546783.2018.1507949>
- Baron J (1993) Why teach Thinking?-An essay. *Appl Psychol* 42(3):191–214. <https://doi.org/10.1111/j.1464-0597.1993.tb00731.x>
- Baron J (2005) *Rationality and intelligence*. Cambridge University Press
- Baron J (2019) Actively open-minded thinking in politics. *Cognition* 188:8–18. <https://doi.org/10.1016/j.cognition.2018.10.004>
- Baron J (2020) Religion, cognitive style, and rational thinking. *Curr Opin Behav Sci* 34:64–68. <https://doi.org/10.1016/j.cobeha.2019.12.015>
- Baron J, Scott S, Fincher K, Emlen Metz S (2015) Why does the cognitive reflection test (sometimes) predict utilitarian moral judgment (and other things)? *J Appl Res Memory Cognition* 4(3):265–284. <https://doi.org/10.1016/j.jarmac.2014.09.003>
- Björklund F, Bäckström M (2008) Individual differences in processing styles: validity of the Rational-Experiential inventory. *Scand J Psychol* 49(5):439–446. <https://doi.org/10.1111/j.1467-9450.2008.00652.x>
- Buckholtz JW, Asplund CL, Dux PE, Zald DH, Gore JC, Jones OD, Marois R (2008) The neural correlates of Third-Party punishment. *Neuron* 60(5):930–940. <https://doi.org/10.1016/j.neuron.2008.10.016>
- Buckholtz J, W, Marois R (2012) The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nat Neurosci* 15(5):Articolo5. <https://doi.org/10.1038/nn.3087>
- Buon M, Jacob P, Loissel E, Dupoux E (2013) A non-mentalistic cause-based heuristic in human social evaluations. *Cognition* 126(2):149–155. <https://doi.org/10.1016/j.cognition.2012.09.006>
- Buon M, Seara-Cardoso A, Viding E (2016) 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(6):1660–1680. <https://doi.org/10.3758/s13423-016-1042-5>
- Burra A, Knobe J (2006) The folk concepts of intention and intentional action: A Cross-Cultural study. *J Cognition Cult* 6(1–2):113–132. <https://doi.org/10.1163/156853706776931222>
- Cushman F (2008) Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition* 108(2):353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Cushman F (2013) Action, outcome, and value: A Dual-System framework for morality. *Personality Social Psychol Rev* 17(3):273–292. <https://doi.org/10.1177/1088868313495594>
- Cushman F (2015) Deconstructing intent to reconstruct morality. *Curr Opin Psychol* 6:97–103. <https://doi.org/10.1016/j.copsyc.2015.06.003>
- Cushman F, Young L, Hauser M (2006) The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. *Psychol Sci* 17(12):1082–1089. <https://doi.org/10.1111/j.1467-9280.2006.01834.x>
- Darley JM (2009) *Morality in the Law: The Psychological Foundations of Citizens' Desires to Punish Transgressions* (SSRN Scholarly Paper 1600001). <https://doi.org/10.1146/annurev.lawsocsci.4.110707.172335>
- De Neys W, Pennycook G (2019) Logic, fast and slow: advances in Dual-Process theorizing. *Curr Dir Psychol Sci* 28(5):503–509. <https://doi.org/10.1177/0963721419855658>
- Decety J, Cacioppo S (2012) The speed of morality: A high-density electrical neuroimaging study. *J Neurophysiol* 108(11):3068–3072. <https://doi.org/10.1152/jn.00473.2012>. Scopus
- Epstein S, Pacini R, Denes-Raj V, Heier H (1996) Individual differences in intuitive-experiential and analytical-rational thinking styles. *J Personal Soc Psychol* 71(2):390. <https://doi.org/10.1037/0022-3514.71.2.390>
- Evans JSBT (2003) In two minds: Dual-process accounts of reasoning. *Trends Cogn Sci* 7(10):454–459. <https://doi.org/10.1016/j.tics.2003.08.012>
- Evans JSBT (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Ann Rev Psychol* 59:255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans J, St. BT, Barston JL, Pollard P (1983) On the conflict between logic and belief in syllogistic reasoning. *Mem Cognit* 11(3):295–306. <https://doi.org/10.3758/BF03196976>
- Evans J, St. BT, Stanovich KE (2013) Dual-Process theories of higher cognition: advancing the debate. *Perspect Psychol Sci* 8(3):223–241. <https://doi.org/10.1177/1745691612460685>
- Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 39(2):175–191. <https://doi.org/10.3758/BF03193146>
- Faul F, Erdfelder E, Buchner A, Lang A-G (2009) Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 41(4):1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Feltz A (2007) The Knobe effect: A brief overview. *J Mind Behav* 28(3/4):265–277
- Flick C, Nuñez N, Laurent SM (2024) If negligence is intentionality's cousin, recklessness is its sibling: differentiating negligence and recklessness from accidents and intentional harm. *J Exp Soc Psychol* 110:104529. <https://doi.org/10.1016/j.jesp.2023.104529>
- Greco E (1993) The emotional Stroop test: A review of the literature. *Psichiatria E Psicoterapia Analitica* 12(3):219–223
- Guglielmo S, Malle BF (2010) Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. *Personality Social Psychol Bulletin* 36(12):1635–1647. <https://doi.org/10.1177/0146167210386733>
- Janssen EM, Verkoeijen PPJL, Heijltjes AEG, Mainhard T, van Peppen LM, van Gog T (2020) Psychometric properties of the actively Open-minded thinking scale. *Think Skills Creativity* 36:100659. <https://doi.org/10.1016/j.tsc.2020.100659>
- Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, p 499

- Klauer KC, Musch J, Naumer B (2000) On belief bias in syllogistic reasoning. *Psychol Rev* 107(4):852–884. <https://doi.org/10.1037/0033-295x.107.4.852>
- Knobe J (2003) Intentional action and side effects in ordinary Language. *Analysis* 63(3):190–194
- Knobe J (2022) Morality and Possibility. In M. Vargas & J. M. Doris (A. c. Di), *The Oxford Handbook of Moral Psychology* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198871712.013.18>
- Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E (2006) Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314(5800):829–832. <https://doi.org/10.1126/science.1129156>
- Krueger F, Hoffman M (2016) The emerging neuroscience of Third-Party punishment. *Trends Neurosci* 39(8):499–501. <https://doi.org/10.1016/j.tins.2016.06.004>
- Leslie AM, Knobe J, Cohen A (2006) Acting intentionally and the Side-Effect effect: theory of Mind and moral judgment. *Psychol Sci* 17(5):421–427. <https://doi.org/10.1111/j.1467-9280.2006.01722.x>
- MacLaren VV, Fugelsang JA, Harrigan KA, Dixon MJ (2012) Effects of impulsivity, reinforcement sensitivity, and cognitive style on pathological gambling symptoms among frequent slot machine players. *Pers Individ Differ* 52(3):390–394. <https://doi.org/10.1016/j.paid.2011.10.044>
- Malle BF, Knobe J (1997) The folk concept of intentionality. *J Exp Soc Psychol* 33(2):101–121. <https://doi.org/10.1006/jesp.1996.1314>
- Margoni F, Surian L (2016) Explaining the U-Shaped development of Intent-Based moral judgments. *Front Psychol* 7. <https://doi.org/10.3389/fpsyg.2016.00219>
- Margoni F, Surian L (2022) Judging accidental harm: due care and foreseeability of side effects. *Curr Psychol* 41(12):8774–8783. <https://doi.org/10.1007/s12144-020-01334-7>
- Margoni F, Geipel J, Hadjichristidis C, Surian L (2018) Moral judgment in old age. *Exp Psychol* 65(2):105–114. <https://doi.org/10.1027/1618-3169/a000395>
- Margoni F, Geipel J, Hadjichristidis C, Surian L (2019) The influence of agents' negligence in shaping younger and older adults' moral judgment. *Cogn Dev* 49:116–126. <https://doi.org/10.1016/j.cogdev.2018.12.002>
- Martin JW, Buon M, Cushman F (2021) The effect of cognitive load on Intent-Based moral judgment. *Cogn Sci* 45(4):e12965. <https://doi.org/10.1111/cogs.12965>
- Morley (née Lambell), Evans NJ, St. J. B. T., Handley SJ (2004) Belief bias and figural bias in syllogistic reasoning. *Q J Experimental Psychol Sect A* 57(4):666–692. <https://doi.org/10.1080/02724980343000440>
- Nadelhoffer T (2004) On praise, side effects, and folk ascriptions of intentionality. *J Theoretical Philosophical Psychol* 24(2):196–213. <https://doi.org/10.1037/h0091241>
- Ngo L, Kelly M, Coutlee CG, Carter RM, Sinnott-Armstrong W, Huettel SA (2015) Two distinct moral mechanisms for ascribing and denying intentionality. *Sci Rep* 5(1):17390. <https://doi.org/10.1038/srep17390>
- Nobes G, Martin JW (2022) They should have known better: the roles of negligence and outcome in moral judgements of accidental actions. *Br J Psychol* 113(2):370–395. <https://doi.org/10.1111/bjop.12536>
- Nobes G, Panagiotaki G, Pawson C (2009) The influence of negligence, intention, and outcome on children's moral judgments. *J Exp Child Psychol* 104(4):382–397. <https://doi.org/10.1016/j.jecp.2009.08.001>
- Nobes G, Panagiotaki G, Engelhardt PE (2017) The development of intention-based morality: the influence of intention salience and recency, negligence, and outcome on children's and adults' judgments. *Dev Psychol* 53(10):1895–1911. <https://doi.org/10.1037/dev0000380>
- Núñez N, Laurent S, Gray JM (2014) Is negligence a first cousin to intentionality? lay conceptions of negligence and its relationship to intentionality. *Applied Cognitive Psychology*, 28(1), 55–65. <https://doi.org/10.1002/acp.2957>
- Ozono H, Sakakibara R (2024) The moderating role of reflective thinking on personal factors affecting belief in conspiracy theories. *Applied Cognitive Psychology*, 38(1), e4142. <https://doi.org/10.1002/acp.4142>
- Ozono H, Sakakibara R (2024) The moderating role of reflective thinking on personal factors affecting belief in conspiracy theories. *Appl Cogn Psychol* 38(1):e4142. <https://doi.org/10.1002/acp.4142>
- Pacini R, Epstein S (1999) The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *J Personal Soc Psychol* 76(6):972–987. <https://doi.org/10.1037/0022-3514.76.6.972>
- Patil I, Trémolière B (2021) Reasoning supports forgiving accidental harms. *Scientific Reports*, 11(1), Article 1. <https://doi.org/10.1038/s41598-021-93908-z>
- Patil I, Zucchelli MM, Kool W, Campbell S, Fornasier F, Calò M, Silani G, Cikara M, Cushman F (2021) Reasoning supports utilitarian resolutions to moral dilemmas across diverse measures. *J Personal Soc Psychol* 120(2):443–460. <https://doi.org/10.1037/pssp0000281.supp>
- Pettit D, Knobe J (2009) The pervasive impact of moral judgment. *Mind Lang* 24(5):586–604. <https://doi.org/10.1111/j.1468-0017.2009.01375.x>
- Pinillos N, Smith N, Nair G, Marchetto P, Mun C (2011) Philosophy's new challenge: experiments and intentional action. *Mind Lang* 26. <https://doi.org/10.1111/j.1468-0017.2010.01412.x>
- Robinson P, Kurzban R, Jones O (2007) The origins of shared intuitions of justice. *VANDERBILT LAW Rev* 60(6):1633–
- Schwartz F, Djeriouat H, Trémolière B (2022) Judging accidental harm: reasoning style modulates the weight of intention and harm severity. *Q J Experimental Psychol* 75(12):2366–2381. <https://doi.org/10.1177/17470218221089964>
- Schwartz F, Passemar A, Djeriouat H, Trémolière B (2024) The Temporal dynamics of third-party moral judgment of harm transgressions: answers from a 2-response paradigm. *Think Reasoning* 30(1):109–134. <https://doi.org/10.1080/13546783.2023.2210847>
- Scopus
- Shen FX, Hoffman MB, Jones OD, Greene JD (2011) Sorting guilty Minds. *New York Univ Law Rev* 86(5):1306–1360
- Sladek RM, Bond MJ, Phillips PA (2010) Age and gender differences in preferences for rational and experiential thinking. *Pers Individ Differ* 49(8):907–911. <https://doi.org/10.1016/j.paid.2010.07.028>
- Sloman SA (1996) The empirical case for two systems of reasoning. *Psychol Bull* 119(1):3–22. <https://doi.org/10.1037/0033-2909.119.1.3>
- Spitzer M, Fischbacher U, Herrnberger B, Grön G, Fehr E (2007) The neural signature of social norm compliance. *Neuron* 56(1):185–196. <https://doi.org/10.1016/j.neuron.2007.09.011>
- Stanovich KE, Toplak ME (2023) Actively Open-Minded thinking and its measurement. *J Intell* 11(2):Article2. <https://doi.org/10.3390/jintelligence11020027>
- Stanovich KE, West RF (1997) Reasoning independently of prior beliefs and individual differences in actively Open-Minded thinking. *J Educ Psychol* 89(2):342–357
- Stanovich KE, West RF (2000) Individual differences in reasoning: implications for the rationality debate? *Behav Brain Sci* 23(5):645–665. <https://doi.org/10.1017/S0140525X00003435>
- Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* 18(6):643–662. <https://doi.org/10.1037/h0054651>

- Svedholm-Häkkinen AM, Lindeman M (2017) Intuitive and deliberative empathizers and systemizers. *J Pers* 85(5):593–602. <https://doi.org/10.1111/jopy.12263>
- Szaszi B, Szollosi A, Palfi B, Aczel B (2017) The cognitive reflection test revisited: exploring the ways individuals solve the test. *Think Reason*. <https://www.tandfonline.com/doi/abs/> <https://doi.org/10.1080/13546783.2017.1292954>
- Thompson V, Evans J, St. BT (2012) Belief bias in informal reasoning. *Think Reason* 18(3):278–310. <https://doi.org/10.1080/13546783.2012.670752>
- Trémolière B, Neys WD, Bonnefon J-F (2017) Reasoning and moral judgment: A common experimental toolbox. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning*. Routledge
- Witteman C, van den Bercken J, Claes L, Godoy A (2009) Assessing rational and intuitive thinking styles. *Eur J Psychol Assess* 25(1):39–47. <https://doi.org/10.1027/1015-5759.25.1.39>
- Yang Q, Hoffman M, Krueger F (2024) The science of justice: the neuropsychology of social punishment. *Neurosci Biobehav Rev* 157:105525. <https://doi.org/10.1016/j.neubiorev.2023.105525>
- Young L, Saxe R (2009) Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia* 47(10):2065–2072. <https://doi.org/10.1016/j.neuropsychologia.2009.03.020>
- Zucchelli MM, Nori R, Gambetti E, Giusberti F (2018) The influence of high autistic personality traits on the attribution of intentionality in typically developing individuals. *J Cogn Psychol* 30(8):840–853. <https://doi.org/10.1080/20445911.2018.1530241>
- Zucchelli MM, Starita F, Bertini C, Giusberti F, Ciaramelli E (2019) Intentionality attribution and emotion: the Knobe effect in alexithymia. *Cognition* 191:103978. <https://doi.org/10.1016/j.cognition.2019.05.015>
- Zucchelli MM, Trotti MAA, Pavan N, Piccardi A, L., Nori R (2025) The dual process model: the effect of cognitive load on the ascription of intentionality. *Front Psychol* 16:1451590. <https://doi.org/10.3389/fpsyg.2025.1451590>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.