# A diffusion model analysis of belief bias: Different cognitive mechanisms explain how cognitive abilities and thinking styles contribute to conflict resolution in reasoning[☆]

Anna-Lena Schubert [a,*], Mário B. Ferreira [b], André Mata [b], Ben Riemenschneider [a]

[a] *Institute of Psychology, Heidelberg University, Heidelberg, Germany*
[b] *CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Portugal*

## ARTICLE INFO

## ABSTRACT

Recent results have challenged the widespread assumption of dual process models of belief bias that sound reasoning relies on slow, careful reflection, whereas biased reasoning is based on fast intuition. Instead, parallel process models of reasoning suggest that rule- and belief-based problem features are processed in parallel and that reasoning problems that elicit a conflict between rule- and belief-based problem features may also elicit more than one Type 1 response. This has important implications for individual-differences research on reasoning, because rule-based responses by certain individuals may reflect that these individuals were either more likely to give a rule-based default response or that they successfully inhibited and overrode a belief-based default response. In two studies, we used the diffusion model to describe decision making in a transitive reasoning task. In Study 1, 41 participants were asked to evaluate conclusions based on their validity. In Study 2, 133 participants evaluated conclusions based on their validity or believability. We tested which diffusion model parameters reflected conflict resolution and related those model parameters to individual differences in cognitive abilities and thinking styles. Individual differences in need for cognition predicted successful conflict resolution under logic instruction, which suggests that a disposition to engage in reflective thinking facilitates the inhibition and override of Type 1 responses. Intelligence, however, was negatively related to successful conflict resolution under belief instruction, which suggests that individuals with high cognitive abilities quickly generated a higher-level logical response that interfered with their ability to evaluate lower-level intrinsic problem features. Taken together, this double dissociation indicates that cognitive abilities and thinking styles affect the processing of conflict information through different mechanisms and at different stages: Greater cognitive abilities facilitate the efficient creation of decoupled problem representations, whereas a greater disposition to engage in critical thinking facilitates the detection and override of Type 1 responses.

Prior beliefs about the world shape our perception of new information. Often, these prior beliefs allow us to make fast decisions based on previous experiences and world knowledge without having to resort to effortful processing, enabling adaptive behavior in a wide range of standard situations. However, prior beliefs may bias our reasoning, resulting in suboptimal decision making in certain situations. Nearly a century's worth of research on the phenomenon of *belief bias* has shown that people are more likely to ignore the logical validity of a conclusion if it is in conflict with their prior beliefs (Evans, Barston, & Pollard, 1983). However, not everyone is equally likely to be afflicted by task-irrelevant prior beliefs: Individuals with high cognitive abilities and a disposition to engage in critical thinking are less prone to fall for belief-based reasoning than others (De Neys, 2006; Stanovich, West, & Toplak, 2016).

The traditional view is that correct responses to problems with a conflict between rule- and belief-based reasoning rely on slow, effortful processing, as opposed to fast and autonomous processing (Evans, 2008). More formally, dual-process model accounts of belief bias distinguish between undemanding, belief-based Type 1 processes and cognitively demanding, rule-based Type 2 processes. It is proposed that

---

Type 1 processing endorses/opposes a conclusion if it is consistent/inconsistent with prior beliefs regardless of its logical validity. If this default response conflicts with a logic-based response, and if that conflict is detected, Type 2 processes may be used to inhibit and replace the initial response.

Individual differences in working memory capacity and the disposition to engage in critical thinking have been shown to predict performance in reasoning tasks with a conflict between rule- and belief-based processing, which has been interpreted as evidence for the successful engagement of deliberate Type 2 processing (De Neys, 2006; Stanovich et al., 2016). Cognitive abilities in general and working memory capacity in particular have been suggested to facilitate the inhibition of premature responses resulting from Type 1 processing (Stanovich, 2009, 2010). Moreover, they have been suggested to be helpful in creating a decoupled representation of the problem, which can then be used to generate an alternative, rule-based response (Stanovich, 2009, 2010). In addition, the disposition and ability to engage in critical thinking has been suggested to facilitate the initiation of decoupled reasoning and the successful override of Type 1 responses (Stanovich & West, 2000; Toplak, West, & Stanovich, 2011). According to Stanovich's theory of rational thinking (Stanovich, 2009, 2010), cognitive abilities (the algorithmic mind) and critical thinking dispositions (the reflective mind) contribute together to the successful override of incorrect responses generated by belief-based Type 1 reasoning (the autonomous mind).

## 1. Challenges to dual-process theories of reasoning

Although there is substantial evidence supporting the key assumptions of dual-process theories (for a review see Ball, Thompson, & Stupple, 2018), several recent findings have challenged the assumption that intuitive Type 1 processes generate an initial belief-based response that may later be inhibited and overridden by deliberate Type 2 processes.

First, belief-based reasoning may not be as fast and autonomous as commonly presumed. Indeed, shorter problem presentation times do not increase belief bias (Evans, Handley, & Bacon, 2009), and granting individuals more time may even increase their propensity to draw belief-based conclusions (Newman, Gibb, & Thompson, 2017). Moreover, when asked to evaluate the believability of a conclusion instead of its validity, participants performed worse when the two sources of information were in conflict (Handley, Newstead, & Trippas, 2011), which cannot be explained by sequential dual-process theories, as relatively slow deliberate processes should not interfere with fast and autonomous belief-based reasoning. These results question the notion of exclusively fast and autonomous Type 1 processes. In addition, performing a concurrent demanding task has been shown to not only impair performance when participants had to evaluate the logical validity of conclusions, but also when they had to evaluate their believability (Howarth, Handley, & Walsh, 2016), suggesting that both judgments rely on working memory resources.

Furthermore, recent findings challenge the assumption that logic-based reasoning is slow and requires effortful processing. Individuals sometimes seem to detect the conflict between logic and believability even when defaulting to the belief-based response, suggesting that people may have logical intuitions (De Neys, 2012, 2014). Moreover, individuals have been shown to be able to evaluate the validity of inferences containing abstract non-words even under challenging time constraints (Newman et al., 2017).

These findings are hard to reconcile with core ideas of traditional dual-process models. Instead, they suggest that rule- and belief-related problem features are processed in parallel and that reasoning problems that elicit a conflict between rule- and belief-based thinking may also elicit more than one Type 1 response (De Neys, 2012, 2014; Handley & Trippas, 2015; Pennycook, Fugelsang, & Koehler, 2015). Problem features such as feature complexity or fluency may then determine whether participants give a rule- or belief-based default response (Handley & Trippas, 2015; Pennycook et al., 2015; Trippas, Thompson, & Handley, 2017). Nevertheless, a divergence between rule- and belief-based processing will always cause mutual interference, resulting in lower accuracies and higher reaction times (Trippas et al., 2017).

## 2. Implications for individual differences research

These challenges to classic dual-process theories have far-reaching implications for the interpretation of results from individual differences research on belief bias. If belief- and rule-based problem features are processed in parallel, rule-based responses by individuals with certain personality traits and thinking dispositions may reflect that these individuals were either more likely to give a rule-based default response or that they successfully inhibited a belief-based default response and responded with a rule-based deliberate response.

In a recent study, Thompson, Pennycook, Trippas, and Evans (2018) found that, in conflict trials (where the logical validity of a conclusion was incongruent with its believability), individuals with higher cognitive abilities performed worse when asked to assess the believability of a conclusion than when asked to assess the logical validity. Conversely, individuals with lower cognitive abilities were better at evaluating the believability of a conclusion than its validity. These results suggest that default responses may differ between high and low ability reasoners, and that at least part of the association between cognitive abilities and rule-based reasoning can be attributed to individual differences in the availability of belief- and rule-based default responses. This notion is further supported by previous research showing that more intelligent individuals not only engage in more Type 2-processing, but are also more likely to give rule-based responses as a first, intuitive answer, as compared to less intelligent individuals (Thompson & Johnson, 2014). Together, these results suggest that individuals with higher cognitive abilities are less afflicted by the belief bias because they are more likely to generate a rule-based default response and are also better at inhibiting and overriding belief-based default responses than individuals with lower cognitive abilities.

Despite these compelling results, which challenge core predictions of sequential dual-process theories of reasoning, some open questions remain. A detailed inspection of the results by Thompson et al. (2018) reveals that only participants falling into the first quartile of the ability distribution showed impairments in their reasoning performance when asked to evaluate conclusions with a conflict between belief- and rule-based responses according to their validity than when they were asked to evaluate these conclusions according to their believability. In comparison, all other cognitive ability groups (i.e., both moderately and highly able participants) were better at assessing the validity of a problem than its believability when there was a conflict between the two sources of information. This suggests that not only individuals with very high cognitive abilities, but almost everyone was better at assessing a problem's logical validity than they were at assessing its believability. An implication of this result would be that the rule-based response was the near-universal dominant default response for everyone except those individuals with very low cognitive abilities. This conclusion is seemingly at odds with the robust phenomenon of belief bias and the deterioration of performance in conflict trials in comparison to trials without a conflict between rule- and belief-based responding (Evans et al., 1983).

However, this result – that only the lowest ability group differed from the rest of the sample in their response behavior – may be an artifact of sample characteristics and the cognitive ability tests employed by Thompson et al. (2018). Because their sample consisted of university students who are likely to show above-average cognitive abilities due to selection processes, the intelligence test used in their study (the Shipley-2; Shipley, Christian, Martin, & Klein, 2009) may not have been difficult enough to differentiate across their sample's whole range of cognitive abilities. In addition, they used a numeracy test that has been shown to

differentiate in highly educated samples but that has not been related to standard intelligence tests (Lipkus, Samsa, & Rimer, 2016). Therefore, it would be important to demonstrate that these results replicate with an intelligence test that is appropriate for highly educated samples and that shows a high loading on general intelligence (*g*). In the present paper, we therefore included Raven's Advances Progressive Matrices (APM; Raven, Court, & Raven, 1994) as an intelligence test, which is suited for measuring intelligence in gifted adolescents and adults and has been shown to have a very high *g*-loading (Gignac, 2015).

Moreover, the present research tests whether the same kind of moderation is found for any kind of cognitive ability measure, or whether different interaction patterns emerge for different cognitive abilities and thinking styles (e.g., fluid intelligence, working memory capacity, critical thinking disposition and ability). Justified by the results from a principal component analysis, Thompson et al. (2018) collapsed all individual differences variables (intelligence, numeracy, reflective thinking, and actively open-minded thinking) into a single cognitive ability variable. This approach leads to an increase in the reliability of the aggregate cognitive ability variable, but comes with a loss of granularity. An inspection of component loadings shows considerable heterogeneity in component loadings with the first principal component being dominated by the intelligence test. This suggests that interactions between cognitive ability variables and performance in the belief bias paradigm may differ substantially across different measures of cognitive ability. In particular, reasoners with high cognitive abilities (i.e., high fluid intelligence and working memory capacity) may be able to quickly create a rule-based default response after reading a problem because they are better at creating higher-level relational structures that temporarily bind elemental perceptual or memorial representations and that can be effectively abstracted from intrinsic features of low-level representations (Chuderski, 2014; Oberauer, Süß, Wilhelm, & Wittmann, 2008). Hence, these reasoners might immediately generate a rule-based response even when asked to evaluate the believability of a problem, due to their fast and efficient processing of higher-order relational structures, subsequently resulting in substantial interference from rule-based Type 1 processes. In comparison, individuals with a disposition to engage in reflective thinking but low cognitive abilities may not generate a dominant rule-based default response as quickly and may therefore experience less interference from rule-based Type 1 processes when asked to assess the believability of a problem. In order to test whether individuals with high cognitive abilities and individuals with a disposition to engage in reflective thinking differ, the moderating effects of both measures on rule- and belief-based responding need to be considered separately. In the present paper, we therefore measured several indicators of cognitive abilities and thinking styles to comprehensively investigate how individual differences in both traits moderate rule- and belief-based reasoning in a transitive reasoning task.

## 3. Performance measures in transitive reasoning tasks

Transitive reasoning is a form of inferential reasoning that subsumes relational inferences made from relations between objects, individuals, or entities (Piaget, 1928; Sternberg, 1980b). If A > B and B > C, it can be inferred that A > C. It plays an important role in normal development and has been demonstrated to underlie many socio-cognitive processes including mathematical reasoning, text processing, trusting others, and developing relationships (Wright, 2001; Wright & Smailes, 2015). Successful transitive reasoning in tasks comparable to the ones typically used in research on belief bias requires verbal and spatial in addition to operational reasoning abilities (Piaget, Inhelder, & Szeminska, 1960; Sternberg, 1980a).

One limitation of previous research on the belief bias is that most studies only considered accuracy as the performance measure in reasoning tasks while largely ignoring response times despite their theoretical importance in distinguishing between fast intuitive and slow deliberate processing (e.g., Evans et al., 2009; Thompson et al., 2018). The studies that did assess response times typically found that they were longer in trials with a conflict between rule- and belief-based responding than in no-conflict trials, although evidence for effects of increased processing demands and its moderation by task instructions was less compelling and more inconsistent for response times than for accuracies (Handley et al., 2011; Thompson, Prowse Turner, & Pennycook, 2011; Trippas et al., 2017). At the same time, several studies that used varying presentation times and response deadlines have contributed to theory development by challenging predictions from classical dual process accounts (e.g., Evans et al., 2009; Newman et al., 2017). Hence, research on the belief bias would likely benefit from simultaneously considering responses and response times in order to assess how much processing time is needed for belief-based and rule-based responding.

One reason why previous research has found rather inconsistent results for response time measures may lie in speed-accuracy tradeoffs that can lead to substantial shifts in response time distributions. Even when instructed to respond as quickly and accurately as possible, individuals still differ in the degree to which they put relative emphasis on speed and accuracy in experimental tasks (Heitz, 2014). These shifts are not easily detectable by visual inspection or traditional methods of statistical inference, because substantial shifts in response time (e.g., an increase of 200 ms) are often accompanied by very subtle changes in accuracy (e.g., an increase of only 3%), which can lead to the erroneous conclusion that there are no speed-accuracy tradeoffs (Forstmann et al., 2011; Pew, 1969). This concern is particularly relevant in individual differences research, where individuals with higher abilities may strategically adjust their decision criteria differently than individuals with lower abilities to adapt to task demands (Draheim, Hicks, & Engle, 2016; Draheim, Mashburn, Martin, & Engle, 2019; Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007). Therefore, it is advisable to explicitly consider speed-accuracy tradeoffs by either using measures that integrate response times and accuracies (Hughes, Linck, Bowles, Koeth, & Bunting, 2014; Liesefeld & Janczyk, 2019; Vandierendonck, 2017), or by accounting for them by using mathematical models of decision making (Brown & Heathcote, 2008; Forstmann et al., 2011; Ratcliff, 1978).

The use of mathematical models of decision making has further benefits, as it transforms observed performance measures into parameters of process models. Non-overlapping interactions between two experimental factors can be removed through nonlinear transformations of the measurement scale (Bamber, 1979; Loftus, 1978), severely limiting the interpretability of these interactions. Therefore, it is crucial to make informed decisions about the measurement scales of cognitive processes, as conclusions may not generalize beyond the specific scales used in data analyses. Because mathematical process models allow to specify transformations between a specific dependent variable and a latent process of interest, using such process models permits to interpret non-overlapping interactions in terms of the modeled latent psychological constructs (Wagenmakers, Krypotos, Criss, & Iverson, 2012).

We therefore analyzed participants' performance by means of diffusion modeling. The diffusion model is a mathematical model of decision making that uses the full distribution of responses and response times to estimate latent model parameters describing the decision process (Ratcliff, 1978; Voss, Nagler, & Lerche, 2013). It has never before, to the best of our knowledge, been used to study belief bias, because typically sparse trial numbers in reasoning tasks complicate model fitting (Trippas et al., 2018). The diffusion model assumes that in two alternative forced-choice problems, evidence is continuously accumulated in a random walk evidence accumulation process until one of two decision thresholds is reached. This evidence accumulation consists of a systematic component, the drift *v*, and normally distributed random noise $s^2$. The path of the diffusion process varies from trial to trial due to random noise, resulting in different process durations and outcomes reflected in response time distributions for the two decision alternatives. Under the assumption that individuals do not show an a-priori bias for

any of the two decision alternatives, the two decision thresholds can be recoded to reflect correct and incorrect responses, which allows to interpret the drift rate $v$ as an ability parameter (Frischkorn & Schubert, 2018; Schubert, Nunez, Hagemann, & Vandekerckhove, 2019; van der Maas, Molenaar, Maris, Kievit, & Borsboom, 2011; Voss et al., 2013; Voss, Voss, & Lerche, 2015).

Most importantly, the model distinguishes between the speed and efficiency of information processing, which are reflected in the strength and direction of the drift rate $v$, and decision criteria, which are reflected in the threshold separation parameter $a$ that quantifies how much evidence needs to be accumulated before making a decision. Together, these two parameters predict the speed and accuracy of decision making, with the drift rate parameter indicating how much evidence is accumulated per unit of time, and the threshold separation parameter indicating how much evidence needs to be accumulated before the evidence accumulation process is terminated. As illustrated in Fig. 1, when the processing of rule- and belief-based problem features converges, drift rates should be large (predicting high accuracies and fast responses), whereas when the processing of rule- and belief-based problem features diverges, drift rates should be small (predicting low accuracies and slow responses). In addition, individuals may adjust their decision criteria and hence their decision thresholds when evaluating problems with a conflict between rule- and belief-based processing or when asked to judge the believability of a problem instead of its logical validity. In the present paper, we therefore used drift rates as an integrative performance measure of both responses and response times that is not affected by speed-accuracy tradeoffs to investigate how individual differences in cognitive abilities and thinking styles moderate rule- and belief-based reasoning.

It has previously been suggested that the diffusion model should only be applied to experimental tasks with mean reaction times not exceeding 1.5 s (Ratcliff & McKoon, 2008), because the underlying model assumptions of single-stage decision making and parameter constancy over time are more likely to be violated in more complex decision tasks. This is relevant for the present paper, because decision making and responding in belief bias tasks may exceed this threshold of 1.5 s. Recent work has shown, however, that the diffusion model can be successfully applied to cognitive tasks with reaction times exceeding 1.5 s and is quite robust against violations of model assumptions (Lerche et al., 2020; Lerche & Voss, 2016, 2019; Ratcliff, 2002; van Ravenzwaaij, Donkin, & Vandekerckhove, 2017). Lerche and Voss (2019) argued that the often-repeated 1.5 s rule is based on an arbitrary value and that while the assumptions of single-stage processing and parameter constancy are more likely to be violated in slower response tasks, multistage evidence accumulation can be accounted for by a chain of diffusion processes that can be approximated by a single diffusion process. They tested the extensibility of the diffusion model to slower RT tasks in three experimental validation studies and found that the results of selective manipulations of model parameters were similar to findings from experimental validation studies in fast RT tasks, which allowed them to conclude that the model can be extended to slower RT tasks. This conclusion is further supported by simulation studies, which demonstrated that the diffusion model still provides a good account of the data

if it is specified in a more parsimonious way than the data-generating model (Lerche & Voss, 2016; Ratcliff, 2002; van Ravenzwaaij et al., 2017). Taken together, these results from simulation and validation studies and the successful application of the diffusion model in studies with more complex decision tasks (e.g., Aschenbrenner, Balota, Gordon, Ratcliff, & Morris, 2016; Lerche et al., 2020; von Krause, Lerche, Schubert, & Voss, 2020) suggest that the model can be applied to describe the conflict between rule- and belief-based reasoning even if reaction times are relatively slow.

The diffusion model is widely popular in individual differences research and has been successfully applied to shed light on individual differences in reasoning (Lerche et al., 2020; Schmiedek et al., 2007; Schmitz & Wilhelm, 2016; Schubert et al., 2019; Schubert, Hagemann, Voss, Schankin, & Bergmann, 2015), affective processing (Vandekerckhove, 2014), and cognitive aging (Ratcliff, Thapar, & McKoon, 2010, 2011; Schubert, Hagemann, Löffler, & Frischkorn, 2020; Schulz-Zhecheva, Voelkle, Beauducel, Biscaldi, & Klein, 2016; von Krause et al., 2020). In particular, several studies have demonstrated that more intelligent individuals show advantages in the speed of evidence accumulation (Lerche et al., 2020; Schmiedek et al., 2007; Schmitz & Wilhelm, 2016; Schubert et al., 2015, 2019), whereas age-related differences in cognitive abilities can be mostly attributed to shifts in decision criteria and the slowing of non-decisional processes (Ratcliff et al., 2010, 2011; Schubert et al., 2020; Schulz-Zhecheva et al., 2016; von Krause et al., 2020). Moreover, psychometric studies have identified which conditions need to be met to allow the reliable assessment of trait-like individual differences in diffusion model parameters (Lerche & Voss, 2017; Schubert, Frischkorn, Hagemann, & Voss, 2016). Taken together, there is a great potential for utilizing diffusion models to shed light on individual differences in reasoning. The present paper builds upon these previous studies to assess which process parameters of rule- and belief-based reasoning are moderated by individual differences in cognitive abilities and thinking styles.

## 4. Study 1

The aim of Study 1 was to demonstrate that the diffusion model can account for the conflict between rule- and belief-based processing and to explore which model parameters relate to individual differences in cognitive abilities and thinking styles. We expected to observe smaller drift rates when the processing of rule- and belief-based problem features diverges, as this pattern of results would reflect a competition between the processing of rule-based and belief-based information as suggested by the parallel dual-process model. Moreover, we expected individuals with greater cognitive abilities (in this first study operationalized in terms of working memory capacity) to show higher drift rates in trials with and without a conflict between rule- and belief-based processing than individuals with lower cognitive abilities, because their greater general efficiency of information-processing should result in faster and more accurate responses regardless of the presence or absence of a conflict between rule- and belief-based problem features. In addition, we expected individuals with greater reflective thinking ability (measured by the cognitive reflection test; Frederick, 2005) and a
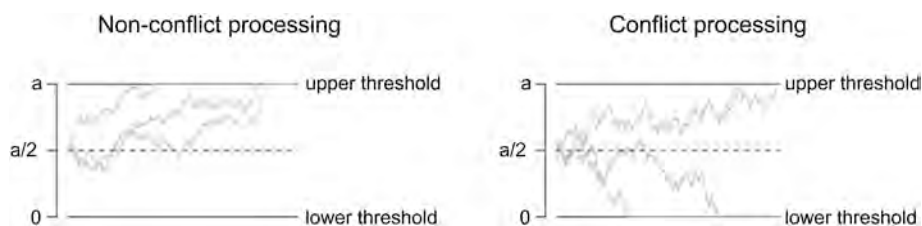


**Fig. 1.** Sample diffusion processes.
*Note.* This figure depicts sample diffusion processes when belief- and rule-based information converge (non-conflict processing) and when belief- and rule-based information diverge (conflict processing) given an invariant boundary separation $a$ across both conditions.

disposition to engage in intellectual activity (measured by the need for cognition scale; Cacioppo & Petty, 1982; Cohen, Stotland, & Wolfe, 1955) to show higher drift rates specifically in trials with a conflict between rule- and belief-based problem features than individuals with lower reflective thinking ability and lower need for cognition, due to their greater ability to disregard belief-based response tendencies that are compelling but incorrect.

### 4.1. Materials and methods

#### 4.1.1. Participants

We recruited $N = 41$ participants (33 females, 8 males) between 18 and 65 years ($M = 26.7$, $SD = 9.0$) with an academic background from the participant pool from the Heidelberg Psychology Department.

#### 4.1.2. Materials

##### 4.1.2.1. Relational reasoning problems.
Participants completed 192 problems (adapted from Banks & Hope, 2014). Each problem consisted of three premises presented sequentially for 3 s each, followed by a conclusion that was either believable or unbelievable. The conclusion was shown for 2 s with the final word omitted, which was subsequently presented until participants gave their response by key-press (see Fig. 2 for details on the task procedure). Following Banks and Hope (2014), we had participants respond to a screen that showed only the final word of the conclusion to ensure that their response times reflected only the speed of evidence accumulation in working memory, but not their reading speed or their willingness to read the premises for a second time. Although this feature of the experimental procedure placed demands on participants' working memory, it did so in a similar way for problems with and without a conflict between rule- and belief-based processing.

The validity of the conclusions was manipulated by changing the premises, whereas the believability of the conclusion was manipulated by reversing the elements in the conclusion. 50% of the trials contained problems with a conflict between rule- and belief-based information, whereas the other 50% of trials were no-conflict problems (see examples in Table 1). Participants completed 12 practice trials with feedback before the start of each of the two experimental blocks. Both mean RTs and accuracies in this task showed acceptable reliabilities, as estimated by Spearman-Brown corrected correlations based on odd-even splits ranging from $r = 0.83$ to $r = 0.95$.

##### 4.1.2.2. Covariates.
We measured working memory capacity (WMC) with the operation span task (Turner & Engle, 1989), following the recommendations for partial scoring by Conway et al. (2005). On average, participants recalled 85% ($SD = 11$%) of the items correctly. In addition, we assessed participants' cognitive and reflective abilities with an extended version of the CRT[1] (Frederick, 2005), which included six additional problems (Shane Frederick, personal correspondence, 02 February 2016). On average, participants solved 4.10 ($SD = 2.63$) problems correctly. We assessed participants' disposition to engage in critical thinking with the need for cognition (NFC) scale of the Rational-Experiential Inventory (REI; Epstein, Pacini, Denes-Raj, & Heier, 1996; Keller, Bohner, & Erb, 2000). Participants' mean NFC score was 5.01 ($SD = 0.68$). Participants also completed the faith in intuition scale of the REI, which is not reported in the present study. All measures showed

moderate-to-good internal consistencies, ranging from $α = 0.78$ for the CRT score over $α = 0.80$ for the NFC scale to $α = 0.87$ for the operation span score.

#### 4.1.3. Procedure

We administered the relational reasoning task followed by the operation span task, the CRT, and the REI. Each session took approximately 1.5 h. While participants completed the relational reasoning task, their EEG was recorded (data not reported here). Participants sat in a dimly-lit, sound-attenuated cabin and were instructed to avoid large movements (e.g., stretching, drinking) except during breaks or inter-trial-intervals.

#### 4.1.4. Data analysis

RTs faster than 100 ms or slower than 5000 ms, or with logarithmized RTs exceeding $\pm 3$ SDs of the mean of each condition were discarded.

A model comparison approach was used to identify which diffusion model parameters differed between experimental conditions in the relational reasoning task. For this purpose, we fit four models to the data: In model a) no diffusion model parameter was allowed to vary across conditions; in b) drift rate was allowed to vary across conditions; in c) boundary separation was allowed to vary across conditions; and in d) both drift rate *and* boundary separation were allowed to vary across conditions. We estimated model parameters with fast-dm-30.2 (Voss & Voss, 2007) using the maximum likelihood approach. In addition to drift rate *v* and boundary separation *a*, we also estimated the non-decision time *t0* and the trial-to-trial variability of the non-decision time $s_{t0}$. The starting point *z* was fixed to *a*/2 and all other model parameters were fixed to zero. These model estimation settings follow recommendations by Lerche and Voss (2016). Model fits were compared based on the Akaike information criterion (AIC; Akaike, 1973) and using Akaike weights (Wagenmakers & Farrell, 2004).

The data supporting the findings of the study are available in the Open Science Framework repository at https://osf.io/ycb2t/ (Schubert, Ferreira, Mata, & Riemenschneider, 2021).

### 4.2. Results

#### 4.2.1. Behavioral data

Table 2 shows mean RTs and accuracies. There was a main effect of conflict on accuracy rates, $F(1,40) = 43.46$, $p < .001$, $ω^2 = 0.34$, such that performance was worse in conflict ($M = 0.58$) than non-conflict ($M = 0.80$) trials. There was no RT difference between conflict and non-conflict trials, $F(1,40) = 0.07$, $p = .790$, $ω^2 = 0.00$.[2]

#### 4.2.2. Diffusion model analysis

Mean AICs were lowest and nearly identical for the model with only drift rate and the model with drift rate and boundary separation allowed to vary across experimental conditions, $AIC_{none} = 575$, $AIC_v = 556$, $AIC_a = 576$, $AIC_{v\&a} = 556$. An inspection of individual Akaike weights showed that the model with only drift rate varying between conditions

---

[1] The current debate regarding the CRT reflects uncertainty concerning its validity: While some argue that the CRT measures individual differences in reflective thinking ability (Frederick, 2005), others argue that it largely measures cognitive and numerical abilities (Blacksmith, Yang, Behrend, & Ruark, 2019; Erceg et al., 2020). Undoubtedly, however, performance in the CRT predicts susceptibility to biases and logical fallacies (e.g., Bialek & Pennycook, 2018; Pennycook & Ross, 2016), making it an important measure to consider in any study on individual differences in bias susceptibility.

[2] To analyze how time-on-task affected the experimental effect, we split trials into four blocks. Accuracy rates increased across the course of the experimental task, $F(3,120) = 3.33$, $p = .022$, $ω^2 = 0.01$, while reaction times decreased, $F(3,120) = 11.74$, $p < .001$, $ω^2 = 0.05$. These training effects were most pronounced from the first to the second block. In addition, we observed an interaction between time-on-task and the congruency condition on reaction times, which reflected that reaction times were faster in the congruent than in the incongruent condition in all blocks but the third, $F(3,120) = 4.76$, $p = .004$, $ω^2 = 0.01$.
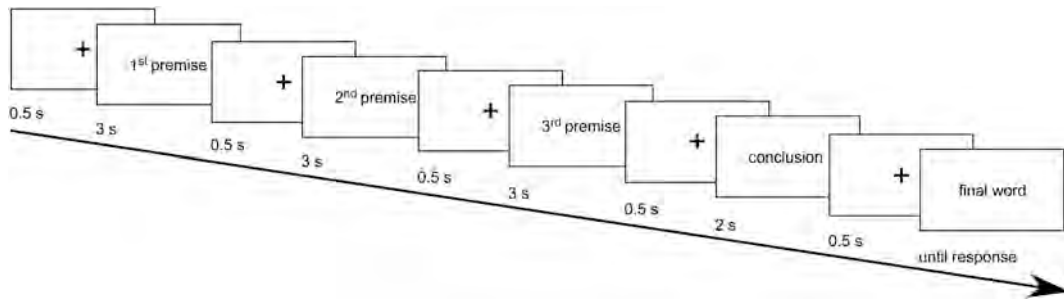
Fig. 2. Procedure of the relational reasoning task.

**Table 1**

Sample items from Banks and Hope (2014) for trials without and with a conflict between rule- and belief-based information.

|  | Example for a relational reasoning task item without a conflict between rule- and belief-based information (non-conflict) | Example for a relational reasoning task item with a conflict between rule- and belief-based information (conflict) |
|---|---|---|
| Premises | cars are faster than skateboards delps are slower than cars delps are faster than bicycles | bicycles are faster than skateboards delps are slower than bicycles delps are faster than cars |
| Conclusion | cars are faster than bicycles | bicycles are faster than cars |

**Table 2**

Mean accuracies (ACC), mean RTs in seconds (RT), mean drift rates (v), mean boundary separations parameters (a), mean non-decision time parameters (t0), and mean inter-trial variabilities of the non-decision times ($s_{t0}$) with SDs in parentheses.

|  | ACC | RT | v | a | t0 | $s_{t0}$ |
|---|---|---|---|---|---|---|
| Conflict | 0.58 (0.20) | 1.33 (0.33) | 0.16 (0.53) | 2.16 (0.38) | 0.38 (0.26) | 0.31 (0.41) |
| Non-conflict | 0.80 (0.10) | 1.32 (0.35) | 0.73 (0.32) | | | |

provided the best account of the data for 49% of the participants, whereas the model with both drift rate and boundary separation varying between conditions provided the best account of the data for 29% of the participants.[3] All subsequent analyses were therefore conducted with parameter estimates from the model in which only drift rates varied between conditions. Statistical analyses confirmed a main effect of conflict on drift rates, $F(1,40) = 32.87$, $p < .001$, $\omega^2 = 0.29$, such that drift rates were lower in conflict ($M = 0.16$) than non-conflict ($M = 0.73$) trials.

We subsequently evaluated the fit of this model using a simulation study and graphical analyses of model fit and found that the diffusion model provided an excellent account of the observed data. There was no evidence that model predictions were systematically biased (for details, see Appendix A).

To evaluate the reliability of model parameters, we conducted odd-even splits of the observed data and estimated model parameters from the resulting data sets using the same model specifications. All model parameters showed acceptable reliabilities, as estimated by Spearman-Brown corrected correlations based on the resulting parameter estimates ranging from $r = 0.85$ to $r = 0.96$. Correlations between model

parameters and observed variables are shown in Table B.1.

### 4.2.3. Moderation by covariates

*4.2.3.1. Accuracy rates.* When individual-difference measures were introduced as continuous covariates into an ANCOVA, the experimental effect on accuracy rates was moderated by individual differences in cognitive reflection and need for cognition, but not working memory capacity. Individuals with better performance in the CRT showed a smaller decrease in accuracy rates than individuals with lower performance in the CRT when there was a conflict between rule- and belief-based processing (see Fig. 3A for an illustration based on a median split), $F(1,39) = 8.95$, $p = .005$, $\omega^2 = 0.10$. We observed the same tendency for individuals high in need for cognition (see Fig. 3B for an illustration based on a median split), who tended to be less affected by conflicts between rule- and belief-based processing than individuals low in need for cognition, $F(1,38) = 3.25$, $p = .079$, $\omega^2 = 0.03$. Working memory capacity, however, did not moderate conflict resolution, $F(1,38) = 0.00$, $p = .987$, $\omega^2 = 0.00$.

*4.2.3.2. Reaction times.* The experimental effect on RTs was not moderated by any of the covariates, all $Fs \leq 0.68$, all $ps \geq 0.144$, all $\omega^2$s = 0.00.

*4.2.3.3. Drift rates.* To assess how the covariates moderated the experimental effect on drift rates, we entered them as continuous variables into an ANCOVA. The experimental effect on drift rates was moderated by individual differences in cognitive reflection and need for cognition, but not working memory capacity. Individuals with better performance in the CRT showed a smaller decrease in drift rates when there was a conflict between rule- and belief-based processing than individuals with lower performance in the CRT (see Fig. 3A for an illustration based on a median split), $F(1,39) = 8.89$, $p = .005$, $\omega^2 = 0.11$.

We observed the same tendency for individuals high in need for cognition, who tended to be less affected by conflicts between rule- and belief-based processing than individuals low in need for cognition (see Fig. 3B for an illustration based on a median split), $F(1,38) = 3.42$, $p = .072$, $\omega^2 = 0.03$. Working memory capacity did not moderate conflict resolution, $F(1,38) = 0.00$, $p = .962$, $\omega^2 = 0.00$.

Instead, individuals with high working memory capacity tended to show high drift rates in both conflict- and non-conflict trials, whereas individuals with greater cognitive reflection and need for cognition showed higher drift rates than individuals with lower cognitive reflection and need for cognition specifically in trials with a conflict between rule- and belief-based processing (see Table 3). None of the individual differences variables was significantly associated with the boundary separation parameter $a$, suggesting that a disposition to engage in critical thinking did not predict greater decision caution.

### 4.3. Discussion

In the first study, we used a diffusion model approach to account for

---

[3] Participants whose data was best described by the model with only drift rate varying between conditions did not differ in any of the covariates from participants whose data was best described by the model with both drift rate and boundary separation varying between conditions, all $Fs \leq 0.43$, all $ps \geq 0.516$, all $\omega^2$s = 0.00.
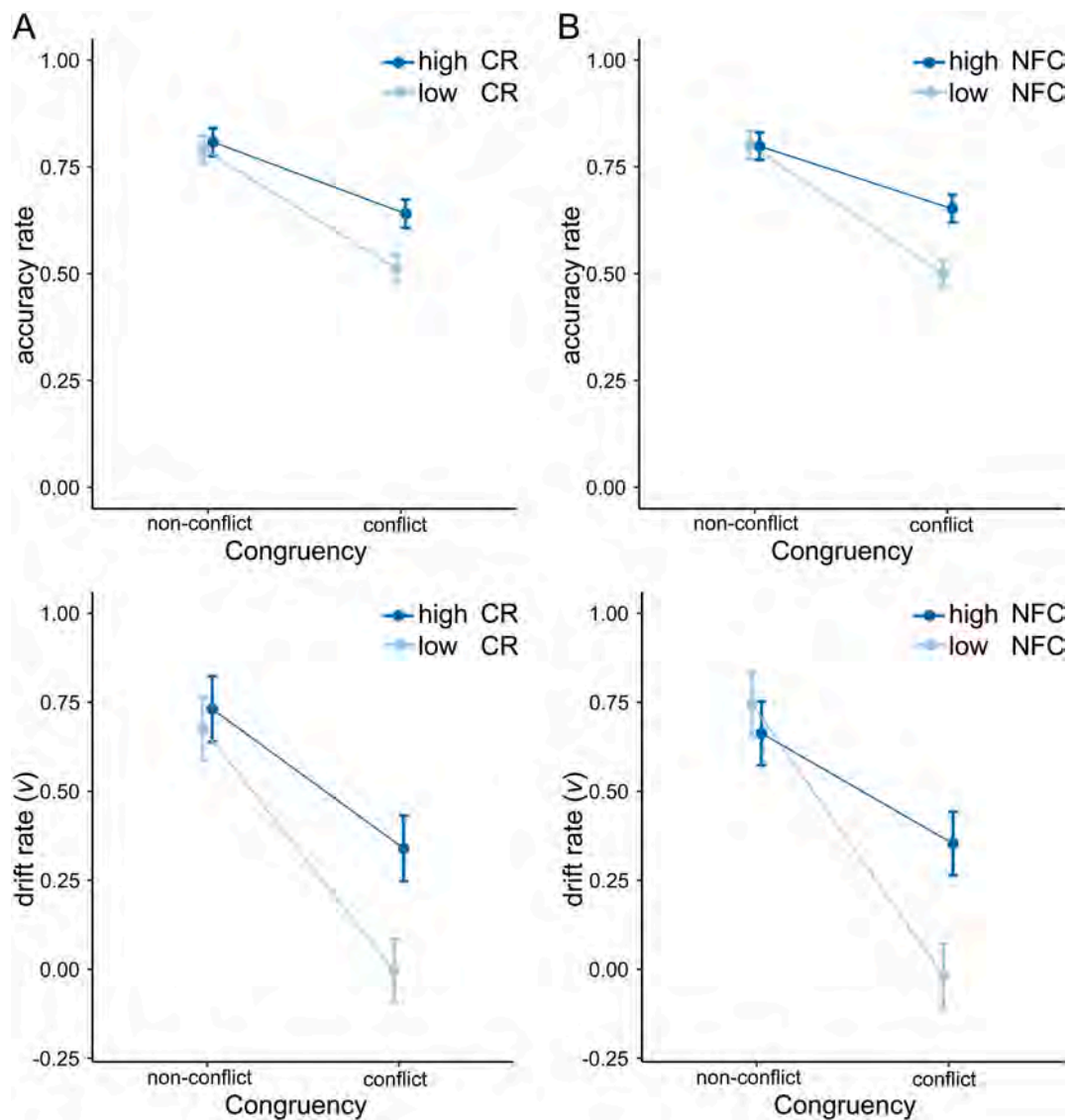
**Fig. 3.** Moderation of the experimental effect by A) cognitive reflection (CR) and B) need for cognition (NFC).
*Note.* CR = cognition reflection; NFC = need for cognition. The median split was only conducted to illustrate the interaction, but that all statistical analyses used continuous covariates.

**Table 3**
Correlations between accuracy rates (ACC), reaction times (RT) and drift rates v in non-conflict- and conflict trials with boundary separation a, working memory capacity (WMC), Cognitive reflection test (CRT) scores, and Need for cognition (NFC) scores.

| | $ACC_{non-conflict}$ | $ACC_{conflict}$ | $ACC_{difference}$ | $RT_{non-conflict}$ | $RT_{conflict}$ | $RT_{difference}$ | $v_{non-conflict}$ | $v_{conflict}$ | $v_{difference}$ | $a$ | WMC | CRT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $ACC_{non-conflict}$ | | | | | | | | | | | | |
| $ACC_{conflict}$ | 0.02 | | | | | | | | | | | |
| $ACC_{difference}$ | 0.43** | −0.89** | | | | | | | | | | |
| $RT_{non-conflict}$ | −0.35** | −0.22 | 0.04 | | | | | | | | | |
| $RT_{conflict}$ | −0.29 | −0.45** | 0.27 | 0.79*** | | | | | | | | |
| $RT_{difference}$ | 0.10 | −0.34* | 0.35* | −0.37* | 0.27 | | | | | | | |
| $v_{non-conflict}$ | 0.87*** | −0.06 | 0.45** | −0.52*** | −0.41** | 0.19 | | | | | | |
| $v_{conflict}$ | 0.01 | 0.98*** | −0.88*** | −0.28 | −0.48** | −0.29 | −0.06 | | | | | |
| $v_{difference}$ | 0.43** | −0.85*** | 0.96*** | −0.03 | 0.19 | 0.34* | 0.55*** | −0.87*** | | | | |
| $a$ | −0.09 | 0.07 | −0.11 | 0.62*** | 0.57*** | −0.12 | −0.41*** | 0.04 | −0.24 | | | |
| WMC | 0.42** | 0.23 | 0.00 | −0.48** | −0.44** | 0.09 | 0.35* | 0.23 | −0.01 | −0.22 | | |
| CRT | 0.22 | 0.59*** | −0.43** | −0.19 | −0.18 | 0.04 | 0.16 | 0.61*** | −0.43** | 0.10 | 0.08 | |
| NFC | 0.11 | 0.37* | −0.28 | 0.01 | −0.09 | −0.13 | 0.04 | 0.37* | −0.29 | 0.12 | −0.05 | 0.46** |

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

the conflict between rule- and belief-based processing. We found that drift rates were smaller when rule- and belief-based problem features conflicted than when they aligned, whereas we found evidence for a criterion shift between conflict and non-conflict problems only in a limited number of participants. This result suggests that the interference caused by a conflict between rule- and belief-based processing impaired evidence accumulation.

The extent to which the conflict between rule- and belief-based processing interfered with evidence accumulation was moderated by individual differences in cognitive abilities and thinking styles. We found dissociations in the way drift rates related to individual differences that may be accounted for in terms of algorithmic and reflective thinking (Stanovich, 2009). While individual differences in reflective thinking (as assessed by the CRT) and the disposition to engage in reflective thinking (as assessed by the NFC scale) were positively related to drift rates specifically in conflict trials, individual differences in algorithmic ability (as assessed by working memory capacity) were related to drift rates in both conflict and non-conflict trials. This association between working memory capacity and overall drift rates is not surprising, as the relational reasoning task placed substantial demands on participants' working memory. In sum, individuals who scored higher on the CRT and who showed greater need for cognition were found to process rule-based information more efficiently and to be less distracted by interfering belief-based information than individuals with lower reflective thinking. It is important to note, however, that these interactions did not overlap and could therefore be removed by transformations of the measurement scale (Bamber, 1979; Loftus, 1978). This is not necessarily problematic, as drift rates are process model parameters that closely reflect the latent cognitive process of interest (Wagenmakers et al., 2012). Nevertheless, it is important to keep in mind that these interactions may still be removed if the data are non-linearly transformed to another measurement scale.

There are three limitations to Study 1. Because the aim of the Study 1 was to demonstrate that the diffusion model can adequately account for the conflict between rule- and belief-based processing, the sample size was comparatively small for investigating how individual differences variables incrementally contribute to transitive reasoning performance. Hence, we could not test whether individual differences in cognitive reflection and need for cognition jointly or independently moderated conflict resolution. Because successfully solving CRT items requires both algorithmic and reflective thinking (Erceg, Galic, & Ružojčić, 2020; Stanovich, 2009, 2010; Toplak et al., 2011), it is possible that CRT solution rates only predicted successful conflict resolution to the degree that they measured reflective thinking ability. However, to test this interpretation of our results, a larger sample size suited for multiple regression analysis would be needed.

Second, we assessed cognitive abilities with a measure of working memory capacity, which does not show as high a loading on general intelligence as proper tests of fluid intelligence (Carroll, 1993).

Third, we only analyzed how individual differences in cognitive abilities and thinking styles moderated performance when participants were instructed to assess the logical validity of a conclusion, but not how individual-difference variables moderated participants' performance in assessing the believability of a conclusion.

In Study 2, we therefore recruited a larger sample of 145 participants who evaluated transitive reasoning problems both on the basis of validity and on the basis of belief, and who in addition to the individual-differences measures included in Study 1 also completed a fluid intelligence test with a high *g*-loading.

## 5. Study 2

When asked to assess a problem based on its validity, we expected cognitive reflection and need for cognition to predict successful conflict resolution (i.e., be associated with the conflict-induced decrease in drift rates), whereas we expected cognitive abilities to be related to the overall rate of evidence accumulation.

When asked to assess a problem based on its believability, however, we expected individuals with greater cognitive abilities to be more affected by interference between rule- and belief-based processing (i.e., to show a larger decrease in drift rates) than individuals with lower cognitive abilities, as was previously found by Thompson et al. (2018). Because Thompson et al. (2018) did not distinguish between cognitive abilities and thinking styles, we made no prediction about a possible dissociation regarding the propensity to be affected by interfering rule-based information when assessing a problem based on belief. We used multivariate analyses to analyze how each individual-differences variable incrementally contributed to transitive reasoning performance.

### 5.1. Materials and methods

#### 5.1.1. Participants

We recruited $N = 145$ participants from the participant pool from the Heidelberg Psychology Department. Of these, four did not complete the experiment. Data from two participants was lost due to technical problems. In addition, six participants were excluded because they took notes during the transitive reasoning task or reported either not understanding the task or forgetting task instructions. The remaining sample consisted of $N = 133$ participants (91 females, 41 males, 1 no information) between 18 and 69 years ($M = 25.1$, $SD = 9.07$). Most participants had an academic background.

#### 5.1.2. Materials

##### 5.1.2.1. Relational reasoning problems.
Participants completed the same 192 reasoning problems as in Study 1. However, in half of the trials, they now had to assess the problems based on the believability of their conclusions, whereas in the other half they had to assess the problems based on the logical validity of their conclusions. This experimental factor ("Instruction") was manipulated block-wise and the order of presentation was counterbalanced across participants. Within each block and identical to Study 1, 50% of the trials contained problems with a conflict between rule- and belief-based information, whereas the other 50% were non-conflict problems. Participants completed six practice trials with feedback before the start of each of the two experimental blocks.

Mean RTs showed good reliabilities, as estimated by Spearman-Brown corrected correlations based on odd-even splits ranging from $r = 0.92$ to $r = 0.97$. In comparison, mean accuracies showed lower reliabilities, with corresponding correlations ranging from $r = 0.55$ to $r = 0.56$ under belief instructions and from $r = 0.81$ to $r = 0.82$ under logic instructions.

##### 5.1.2.2. Covariates.
To assess individual differences in cognitive abilities, we measured working memory capacity with the operation span task (Turner & Engle, 1989) following the recommendations for partial scoring by Conway et al. (2005) and fluid intelligence with an 18-item version of Raven's Advanced Progressive Matrices (APM; Raven et al., 1994). On average, participants recalled 80% ($SD = 11\%$) of the items in the operation span and solved 11.01 ($SD = 2.27$) items of the APM correctly. Cognitive reflection was assessed with an extended version of the CRT that included four additional problems (Frederick, 2005; Toplak, West, & Stanovich, 2014) and need for cognition with the NFC scale of the Rational-Experiential Inventory (REI; Epstein et al., 1996; Keller et al., 2000). Participants also completed the faith in intuition scale of the REI, which is not reported in the present study. On average, participants solved 4.21 ($SD = 2.15$) of the CRT problems correctly and had a NFC score of 2.94 ($SD = 0.94$). All measures showed moderate-to-good internal consistencies, ranging from $\alpha = 0.75$ for the CRT score over $\alpha = 0.78$ for the APM score to $\alpha = 0.88$ for the NFC scale and $\alpha = 0.88$ for the operation span score.

### 5.1.3. Procedure

We administered the relational reasoning task followed by the CRT, the REI, the operation span task, and the APM. Each session took approximately two hours.

### 5.1.4. Data analysis

RTs faster than 100 ms or slower than 5000 ms, or with logarithmized RTs exceeding $\pm 3$ SDs of the mean of each condition, were discarded.

A model comparison approach was used to identify which diffusion model parameters differed between instruction conditions in the relational reasoning task. For this purpose, we estimated different models with either no parameter, drift rate, boundary separation, non-decision time, or any combination of the three parameters varying across instruction conditions. In addition, we allowed the drift rate to vary across congruency conditions in all estimated models based on the results from Study 1. Hence, we estimated a total of eight models. We estimated diffusion model parameters as in Study 1. Model fits were compared based on the Akaike information criterion (AIC; Akaike, 1973) and using Akaike weights (Wagenmakers & Farrell, 2004).

The data supporting the findings of the study are available in the Open Science Framework repository at https://osf.io/ycb2t/ (Schubert et al., 2021).

### 5.2. Results

#### 5.2.1. Behavioral data

Table 4 shows mean RTs and accuracies. There was a main effect of conflict on accuracy rates, $F(1,132) = 249.40$, $p < .001$, $\omega^2 = 0.46$, which were lower in conflict ($M = 0.71$) than in non-conflict ($M = 0.88$) trials. In addition, there was a main effect of instruction on accuracy rates, $F(1,132) = 166.14$, $p < .001$, $\omega^2 = 0.25$, which were lower in the validity ($M = 0.75$) than in believability ($M = 0.84$) condition. Moreover, we observed a significant interaction between the factors Congruency and Instruction, $F(1,132) = 23.45$, $p < .001$, $\omega^2 = 0.04$, such that the detrimental effect of the conflict between rule- and belief-based information was amplified when participants evaluated problems based on their logical validity ($\Delta ACC = 0.20$) compared to when they evaluated problems based on their believability ($\Delta ACC = 0.13$).

Once again, there was no main effect of conflict on RTs, $F(1,132) = 0.05$, $p = .833$, $\omega^2 = 0.00$. We observed a significant main effect of instruction on RTs, $F(1,132) = 45.27$, $p < .001$, $\omega^2 = 0.07$, which were slower in the validity ($M = 1.37$) than in believability ($M = 1.07$) condition. We observed no significant interaction between the two experimental factors, $F(1,132) = 3.00$, $p = .086$, $\omega^2 = 0.00$.[4]

#### 5.2.2. Diffusion model analysis

Mean AICs were lowest for the model with drift rate varying across Congruency conditions (conflict vs. non-conflict) and all three diffusion model parameters (i.e., drift rate, boundary separation, and non-decision time) varying across Instruction conditions (see Table 5). This model provided the best account of the data for 33.83% of the sample, whereas for 25.56% of the sample, varying instructions only had an effect on drift rate and boundary separation, but not on non-decision time.[5] All subsequent analyses were therefore conducted with parameter estimates from the model in which only drift rates varied between congruency conditions and all three parameters varied between instruction conditions.

We subsequently evaluated the fit of this model using a simulation study and graphical analyses of model fit and found that the diffusion model provided a good account of the observed data (for details, see Appendix A). Only accuracies in the congruent condition under belief instructions could not be predicted with a high degree of precision, probably because error rates were extremely low in this condition. There was no evidence that model predictions were systematically biased.

To evaluate the reliability of model parameters, we conducted odd-even splits of the observed data and estimated model parameters from the resulting data sets using the same model specifications. Most parameters showed acceptable reliabilities, as estimated by Spearman-Brown corrected correlations based on the resulting parameter estimates ranging from $r = 0.69$ to $r = 0.92$. Lowest estimates of reliability were observed for drift rates under belief instructions (0.69 and 0.74), which may again reflect the low variability in accuracies in this condition, and the $s_{t0}$ parameter (0.76). Correlations between model parameters and observed variables are shown in Table B.2.

There was a main effect of conflict on drift rates, $F(1,132) = 122.44$, $p < .001$, $\omega^2 = 0.11$, such that drift rates were lower in conflict ($M = 0.71$) than non-conflict ($M = 0.95$) trials. In addition, there was a main effect of instruction on drift rates, $F(1,132) = 311.25$, $p < .001$, $\omega^2 = 0.44$, which were lower in the validity ($M = 0.48$) than in the believability ($M = 1.18$) condition. We also found a small but significant interaction between Congruency and Instruction, $F(1,132) = 6.51$, $p = .012$, $\omega^2 = 0.01$, suggesting that conflict between rule- and belief-based processing led to a greater decrease in drift rates when participants evaluated problems based on their logical validity ($\Delta v = 0.30$) than based on their believability ($\Delta v = 0.18$). Analyses also confirmed a main effect of instruction on boundary separations, $F(1,132) = 20.43$, $p < .001$, which were higher in the validity ($M = 2.26$) than in believability ($M = 2.02$) condition. Finally, analyses failed to confirm a main effect of instruction on non-decision times, $F(1,132) = 0.12$, $p = .730$, $\omega^2 = 0.00$, suggesting that the effect of instruction on non-decision times (i.e., encoding and motor response speed) was very nuanced and only held for a certain part of the sample.

#### 5.2.3. Moderation by covariates

To investigate how individual differences in cognitive abilities and thinking styles moderated the processing of incongruent information under logic- and belief-instruction, we introduced working memory capacity, fluid intelligence, cognitive reflection, and need for cognition as continuous covariates. Because 19 participants had missing values in

---

[4] Again, we analyzed how time-on-task affected the two experimental effects Congruency and Instruction by splitting trials into four blocks. Reaction times decreased linearly across the course of the experimental task, $F(3,393) = 58.49$, $p < .001$, $\omega^2 = 0.04$, and we observed a corresponding increase in accuracy rates that was most pronounced from the first to the second block, $F(3,396) = 26.38$, $p < .001$, $\omega^2 = 0.06$. Participants' performance in trials with a conflict between rule- and belief-based reasoning got better over time, which resulted in a decrease of the estimated effect size of the two-way interaction between Congruency and Instruction on accuracy rates across the four blocks (block one: $\omega^2 = 0.11$, block two: $\omega^2 = 0.04$, block three: $\omega^2 = 0.00$, block four: $\omega^2 = 0.01$), $F(3,396) = 7.55$, $p < .001$, $\omega^2 = 0.02$. Pairwise comparisons supported this interpretation, as the experimental effect of Congruency remained relatively stable over time in the belief condition (block one: $d = 0.62$, block two: $d = 0.89$, block three: $d = 0.60$, block four: $d = 0.62$), while it decreased over time in the logic condition (block one: $d = 1.34$, block two: $d = 1.34$, block three: $d = 0.75$, block four: $d = 0.85$). Overall, this pattern of results suggests that participants benefitted from training effects during the course of the experimental task that were most pronounced when they had to evaluate problems according to their logical validity.

[5] Participants whose data were best described by the model with only drift rate and boundary separation varying between instruction conditions did not significantly differ in any of the covariates from participants whose data were best described by the model with all three parameters varying between instruction conditions, all $Fs \le 1.68$, all $ps \ge 0.200$, all $\omega^2$s $\le 0.01$.

**Table 4**

Mean accuracies (ACC), mean RTs in seconds (RT), mean drift rates (v), mean boundary separations parameters (a), mean non-decision time parameters (t0), and mean inter-trial variabilities of the non-decision times ($s_{t0}$) with SDs in parentheses.

| Instruction | Congruency | ACC | RT | v | a | t0 | $s_{t0}$ |
|---|---|---|---|---|---|---|---|
| Logic | | | | | | | |
| | Conflict | 0.65 (0.15) | 1.36 (0.63) | 0.33 (0.44) | | | |
| | Non-conflict | 0.85 (0.07) | 1.38 (0.58) | 0.63 (0.43) | 2.26 (0.59) | 0.33 (0.37) | |
| Belief | | | | | | | |
| | Conflict | 0.77 (0.13) | 1.08 (0.52) | 1.09 (0.41) | | | 0.28 (0.56) |
| | Non-conflict | 0.90 (0.05) | 1.05 (0.52) | 1.27 (0.45) | 2.02 (0.55) | 0.34 (0.37) | |

**Table 5**

Mean Akaike information criterion (AIC) for each specified model and percentage of the sample whose data were best described by the model. v = drift rate; a = boundary separation, $t_0$ = non-decision time.

| Model parameter varying across experimental conditions | | mean AIC | Best-fitting model for % of participants |
|---|---|---|---|
| Congruency | Instruction | | |
| v | none | 504 | 2.26% |
| v | v | 482 | 14.29% |
| v | a | 482 | 3.76% |
| v | $t_0$ | 494 | 3.01% |
| v | v & a | 468 | 25.56% |
| v | v & $t_0$ | 473 | 13.53% |
| v | a & $t_0$ | 478 | 3.76% |
| v | v & a & $t_0$ | 464 | 33.83% |

at least one of the covariates,[6] they were excluded from the following analyses. Four more participants were identified as multivariate outliers based on their Mahalanobis distance and removed from the following analyses.[7]

*5.2.3.1. Accuracy rates.* We simultaneously introduced all four continuous between-subject covariates into a 2 × 2 ANCOVA with the within-subject factors Congruency and Instruction. After the four continuous covariates were introduced, we observed a significant three-way-interaction between need for cognition, Congruency, and Instruction, $F(1,105) = 6.05, p = .016, \omega^2 = 0.02$. In addition, we observed a significant two-way interaction between fluid intelligence and Instruction, $F(1,105) = 21.41, p < .001, \omega^2 = 0.09$, and between need for cognition and Congruency, $F(1,105) = 4.50, p = .036, \omega^2 = 0.01$.

---

[6] One participant did not complete the NFC scale, performed below guessing probability in the APM, correctly recalled no more than 10% of the memory items in the operation span task, and responded correctly to fewer than 85% of the processing items in the operation span task. Four participants scored below guessing probability in the APM and responded correctly to fewer than 85% of the processing items in the operation span task. Four more participants scored below guessing probability in the APM, one other participant recalled no more than 10% of the memory items in the operation span task, and nine other participants responded correctly to fewer than 85% of the processing items in the operation span task.

[7] The remaining 110 participants showed comparable effects of Congruency and Instruction on behavioral data as the complete sample, with a significant main effect of Congruency on accuracy rates, $F(1,109) = 144.06, p < .001, \omega^2 = 0.29$, but not RTs, $F(1,109) = 0.13, p = .721, \omega^2 = 0.00$, a significant main effect of Instruction on both accuracy rates, $F(1,109) = 203.73, p < .001, \omega^2 = 0.49$, and RTs, $F(1,109) = 41.86, p < .001, \omega^2 = 0.09$, and a significant interaction effect on accuracies, $F(1,109) = 22.27, p < .001, \omega^2 = 0.06$, and RTs, $F(1,109) = 6.97, p = .010, \omega^2 = 0.00$. They also showed comparable effects of Congruency and Instruction on drift rates as the complete sample, with a significant main effect of Congruency, $F(1,109) = 112.23, p < .001, \omega^2 = 0.13$, a significant main effect of Instruction, $F(1,109) = 291.82, p < .001, \omega^2 = 0.48$, and a significant interaction between the two factors, $F(1,109) = 8.51, p = .004, \omega^2 = 0.01$.

There were no further interactions between any of the covariates and any of the experimental effects, all $Fs \leq 3.57$, all $ps \geq 0.062$, all $\omega^2s \leq 0.01$. In addition, we found that more intelligent participants showed overall higher accuracy rates than less intelligent ones independent of experimental conditions, $F(1,105) = 14.67, p < .001, \omega^2 = 0.11$.

To further inspect how participants' intelligence and need for cognition moderated the experimental effects, we calculated difference scores (ΔACC) between the accuracy rates in conflict and non-conflict trials separately for both instruction conditions (see Table 6 for correlations between these difference scores and all covariates). Larger difference scores indicate a greater decrease in accuracies (i.e., greater impairment) when there was a conflict between rule- and belief-based problem features.

Participants' intelligence did not predict how well they processed interfering belief-based information (see Fig. 4A for an illustration based on a median split, where a smaller ΔACC indicates a smaller decrease in accuracy rates), $r = -0.09, p = .376$. Participants with a higher need for cognition, however, showed a smaller difference in accuracy rates when evaluating problems based on their logical validity than individuals with a lower need for cognition, $r = -0.20, p = .033$. This smaller difference indicated that their performance was less impaired when there was a conflict between rule- and belief-based problem features and that they were less affected by interfering belief-based information than individuals with a lower need for cognition (see Fig. 4B). The difference in correlations $\Delta r$ between need for cognition and intelligence was not significant, $\Delta r = 0.11, z = 0.75, p = .227$.

None of the covariates predicted how interfering rule-based information affected participants when they were evaluating problems based on their believability, but a marginally significant correlation between participants' intelligence and the difference in accuracy rates supported the notion that more intelligent individuals may have been somewhat more affected by interfering rule-based information than less intelligent ones, $r = 0.18, p = .060$.

*5.2.3.2. Reaction times.* None of the covariates moderated any of the experimental effects (see Table 6), all $Fs \leq 1.78$, all $ps \geq 0.185$, all $\omega^2s = 0.00$.

*5.2.3.3. Drift rates.* We observed a significant three-way-interaction between fluid intelligence, Congruency, and Instruction (see Fig. 5A for an illustration based on a median split), $F(1,105) = 4.11, p = .045, \omega^2 = 0.00$, and a significant three-way-interaction between need for cognition, Congruency, and Instruction (see Fig. 5B), $F(1,105) = 9.48, p = .003, \omega^2 = 0.01$. Working memory capacity and cognitive reflection did not moderate the experimental effects, all $Fs \leq 0.52$, all $ps \geq 0.474$, all $\omega^2s = 0.00$.

There were no two-way-interactions between any of the covariates and any of the experimental effects, all $Fs \leq 3.58$, all $ps \geq 0.061$, all $\omega^2s \leq 0.01$. In addition, we found a main effect of intelligence, $F(1,105) = 11.31, p = .001, \omega^2 = 0.09$, which suggests that participants with greater cognitive abilities showed a greater general efficiency of information processing as reflected in higher drift rates in both conflict and non-conflict trials than participants with lower cognitive abilities.

To further inspect how participants' intelligence and need for

**Table 6**

Correlations between the difference in accuracies ΔACC, reaction times ΔRT and drift rates Δv between congruency conditions, Advanced Progressive Matrices (APM) scores, working memory capacity (WMC), Cognitive reflection test (CRT) scores, and Need for cognition (NFC) scores.

| | ΔACC | | ΔRT | | Δv | | APM | WMC | CRT |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | logic instruction | belief instruction | logic instruction | belief instruction | logic instruction | belief instruction | | | |
| ΔACC logic instruction | | | | | | | | | |
| ΔACC belief instruction | 0.06 | | | | | | | | |
| ΔRT logic instruction | −0.07 | 0.04 | | | | | | | |
| ΔRT belief instruction | −0.14 | −0.13 | −0.13 | | | | | | |
| Δv logic instruction | 0.86*** | 0.00 | −0.19 | −0.05 | | | | | |
| Δv belief instruction | 0.07 | 0.56*** | 0.14 | −0.52*** | 0.05 | | | | |
| APM | −0.09 | 0.18 | 0.09 | −0.05 | 0.00 | 0.19* | | | |
| WMC | −0.02 | 0.00 | 0.01 | −0.15 | 0.02 | 0.10 | 0.15 | | |
| CRT | −0.03 | −0.07 | 0.07 | 0.07 | 0.02 | −0.07 | 0.38*** | 0.08 | |
| NFC | −0.20* | 0.02 | −0.14 | 0.06 | −0.26** | 0.09 | −0.20* | −0.13 | −0.24* |

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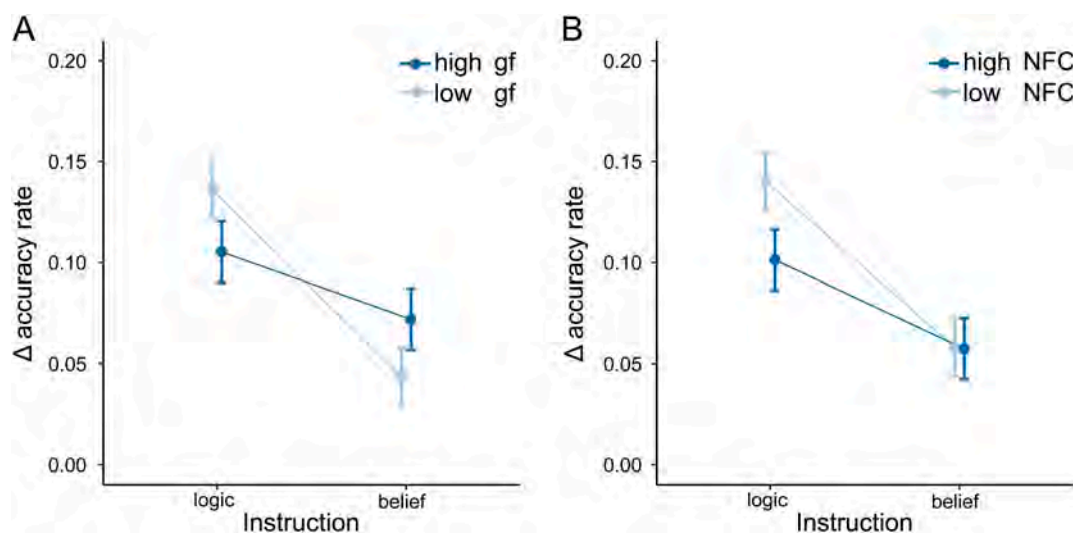


**Fig. 4.** Moderation of the average difference in accuracy rates between non-conflict and conflict trials by Instruction condition and A) fluid intelligence (gf) and B) need for cognition (NFC).
*Note.* The median split was only conducted to illustrate the interaction, but that all statistical analyses used continuous covariates.

cognition moderated the experimental effects, we calculated difference scores Δv between the drift rate in conflict and non-conflict trials separately for both instruction conditions (see Table 6 for correlations between these difference scores and all covariates). Larger difference scores indicated a greater decrease in drift rates (i.e., greater impairment) when there was a conflict between rule- and belief-based problem features than when there was no conflict.

More intelligent participants showed a greater decrease in drift rates than less intelligent ones when asked to assess the believability of a conclusion (see Fig. 5A for an illustration based on a median split), $r = 0.19$, $p = .044$. This greater decrease indicated that their performance was more impaired when there was a conflict between rule- and belief-based problem features and that they were more affected by interfering rule-based information than less intelligent reasoners. When asked to assess the logical validity of a problem, however, more intelligent participants did not differ from less intelligent ones in the way that conflict between rule- and belief-based information affected their performance, $r = 0.00$, $p = .985$. These results suggest that more intelligent participants were more strongly affected by interfering rule-based information than less intelligent ones when evaluating the believability of a problem.

We observed the reverse effect for participants with a high need for cognition, who showed a smaller difference in drift rates than participants with a low need for cognition when asked to assess the logical

validity of a problem, $r = -0.26$, $p = .006$. This smaller difference indicated that their performance was less impaired when there was a conflict between rule- and belief-based problem features and that they were less affected by interfering belief-based information than individuals with a low need for cognition (see Fig. 5B for an illustration based on a median split). When asked to assess a problem based on its believability, however, participants with a high need for cognition did not differ from participants with a low need for cognition in the way that a conflict between rule- and belief-based information affected their performance, $r = 0.09$, $p = .359$. These results suggest that participants with a higher need for cognition were less strongly affected by interfering belief-based information than participants with a lower need for cognition when evaluating the logical validity of a problem.

Finally, we tested if the data contained sufficient evidence to support a full double dissociation. The correlation between Δv and intelligence was not significantly larger than the correlation between Δv and need for cognition when participants were asked to evaluate problems based on their believability, $\Delta r = 0.09$, $z = 0.71$, $p = .238$. However, the correlation between Δv and need for cognition was larger than the correlation between Δv and intelligence when participants were asked to evaluate problems based on their logical validity, $\Delta r = -0.26$, $z = -1.77$, $p = .039$. Taken together, these results show that participants' need for cognition was more predictive of successfully ignoring
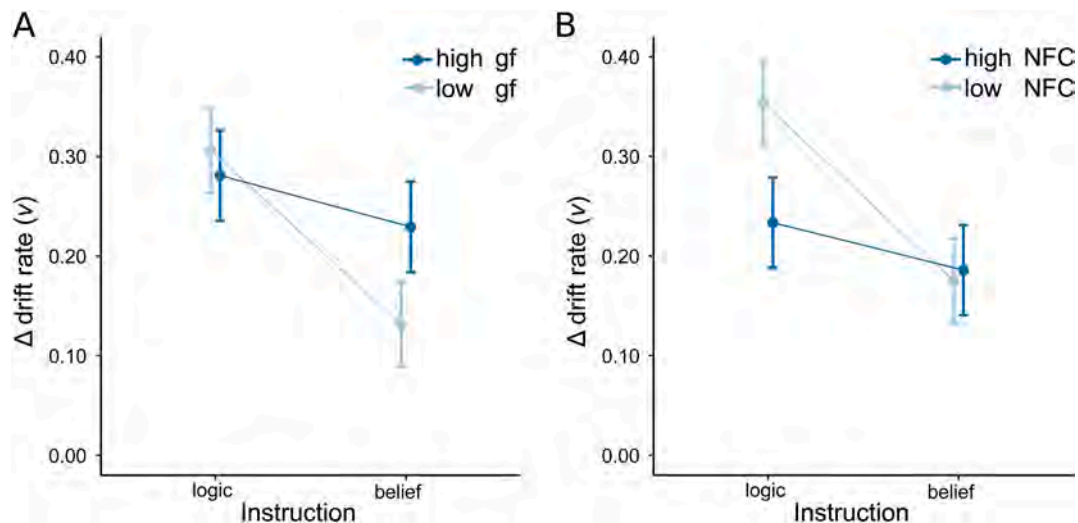
**Fig. 5.** Moderation of the average difference in drift rates (*v*) between non-conflict and conflict trials by Instruction condition and A) fluid intelligence (gf) and B) need for cognition (NFC).

*Note.* Δ*v* = difference in drift rates between congruency conditions; gf = fluid intelligence; NFC = need for cognition. A greater Δ*v* indicates a greater decrease in drift rates when there was a conflict between rule- and belief-based problem features. Note that the median split was only conducted to illustrate the interaction, but that all statistical analyses used continuous covariates.

interfering belief-based information than their intelligence. However, they fail to conclusively demonstrate that participants' intelligence was more predictive of difficulties ignoring interfering rule-based information than their need for cognition.

### 5.3. Discussion

Consistent with results from the first study, we found that participants with a greater need for cognition were less affected by interference between rule- and belief-based processing (i.e., showed a smaller decrease in drift rates) than participants with a lower need for cognition when assessing problems based on their logical validity. When assessing problems based on their believability, however, we found that more intelligent participants were more affected by interference between rule- and belief-based processing (i.e., showed a higher decrease in drift rates) than less intelligent participants. These results tentatively point to a double dissociation in the way that individual differences in fluid intelligence and thinking styles moderate the processing of conflicting rule- and belief-based information. This double dissociation would suggest that different cognitive mechanisms may mediate the associations of fluid intelligence and need for cognition with transitive reasoning performance.

#### 5.3.1. Intelligence and working memory capacity

Individuals with high cognitive abilities seemed to generate a rule-based default response, which then interfered with their ability to evaluate conclusions based on believability, while less intelligent individuals seemed to generate no or weaker logical intuitions. This conclusion is in line with previous findings which demonstrated that more intelligent individuals were more likely to give rule-based responses as a first, intuitive answer than less intelligent individuals (Thompson & Johnson, 2014). Moreover, it is supported by a recent study from Thompson et al. (2018), which also found that individuals with high cognitive abilities performed worse when asked to assess the believability of a conclusion than when asked to assess its logical validity.

Process theories of intelligence can shed some light on why more intelligent reasoners may be more likely to generate a rule-based default response than less intelligent reasoners. Smarter reasoners may more readily and quickly generate a rule-based response, because they are

better at creating higher-level relational structures that temporarily bind elemental perceptual or memorial representations and that can be effectively abstracted from intrinsic features of low-level representations (Chuderski, 2014; Oberauer et al., 2008). Hence, their greater ability to generate abstract relational bindings may actually bias them against evaluating conclusions based on intrinsic and concrete problem features. Surprisingly, however, more intelligent reasoners did not show a corresponding advantage at resolving the conflict between rule- and belief-based information when assessing problems based on their logical validity. Instead, they only showed an overall higher drift rate irrespective of congruency condition, replicating our results from Study 1 and suggesting that individuals with high cognitive abilities benefit from a high general efficiency of information processing as reflected in drift rates in both conflict and non-conflict trials.

It is also important to note that the moderating effect of cognitive abilities on resolving the conflict between rule- and belief-based information only emerged for our fluid intelligence measure, but not for our working memory capacity measure. Although both measures predicted overall drift rates irrespective of congruency and instruction condition, only fluid intelligence moderated the experimental effects. This is not surprising, as the two measures (APM test scores and operation span scores) have been shown to exhibit fundamentally different loadings on a general factor of intelligence (Carroll, 1993) and were also only weakly correlated in the present study, $r = 0.15$. While working memory capacity may sometimes be isomorphic with fluid intelligence (Chuderski, 2013; Kyllonen & Christal, 1990), substantial correlations between working memory capacity and fluid intelligence are typically only found on the latent construct level (Chuderski, 2013; Conway et al., 2005), as the validity of single working memory measures depends on task properties (Conway et al., 2005; Turner & Engle, 1989; Wilhelm, Hildebrandt, & Oberauer, 2013). Hence, our results only highlight that individual differences in cognitive abilities either need to be assessed with a comprehensive test battery of working memory tasks that allows latent variable modeling or with a cognitive ability test that shows a sufficiently high *g*-loading to allow economic assessment.

#### 5.3.2. Need for cognition and cognitive reflection

Need for cognition predicted successful conflict resolution when evaluating conclusions based on their logical validity. In addition, participants with a high need for cognition did not show a bias towards rule-

based responding when evaluating conclusions based on their believability, which suggests that different cognitive mechanisms may account for the associations of rule-based reasoning with cognitive abilities and reflective thinking styles, respectively. Because participants with a high need for cognition had no disadvantage in evaluating conflict problems based on their believability, it is unlikely that they intuitively generated a dominant rule-based default response. Instead, our results suggest that individuals with a disposition to engage in reflective thinking may be as likely to generate a dominant belief-based default response as other individuals, but may be better at successfully inhibiting and overriding the incorrect belief-based response.

Again, it should be noted that the moderating effect of reflective thinking only emerged for need for cognition as a measure of dispositional reflective thinking, but not for cognitive reflection test performance as a measure of reflective thinking ability. We also found that CRT performance was more strongly related to intelligence test performance, $r = 0.38$, than to need for cognition, $r = -0.24$. This finding is not a fluke in our data, but rather it is consistent with a recent study that systematically evaluated the validity of two CRT versions and concluded that they did not access the construct of cognitive reflection, but rather individual differences in intelligence and numerical abilities (Erceg et al., 2020). This conclusion is also supported by a recent study by Chuderski and Jastrzębski (2018), who found strong relations between participants' performance in CRT items and their fluid intelligence. Moreover, Erceg et al. (2020) also failed to find any association between CRT scores and reasoning performance in a belief bias task once individual differences in intelligence and numerical abilities were accounted for. In addition, recent findings have indicated that the CRT may only measure individual differences in cognitive reflection in less intelligent individuals who do not possess strong logical intuitions, but that it is not an accurate measure of cognitive reflection in more intelligent individuals with strong numerical abilities (Erceg, Galic, & Bubić, 2019). Hence, our findings chime in with other studies questioning the validity of the CRT as a measure of reflective thinking. In light of this, it is surprising that we found evidence for convergent validity between the CRT and NFC in Study 1, $r = 0.46$. This discrepancy between studies may reflect differences in the sample composition and the specific CRT items included in the two studies.[8] Because we used a published extension of the CRT in the second study, whereas we included items only conveyed by personal communication in the first one, and because participant numbers were much larger and the sample composition more diverse in the second study than the first, we believe that the data from the second study contain more compelling evidence than the data from the first study, underscoring ours and others' notion that the CRT largely measures cognitive abilities (Chuderski & Jastrzębski, 2018; Erceg et al., 2019, 2020).

## 6. General discussion

In two studies, we used stochastic diffusion models to analyze how individuals processed information during a transitive reasoning task when the logical validity of a conclusion conflicted with its believability. We found that the belief bias effect could be mapped onto the drift rate parameter of the diffusion model, which describes the speed and efficiency of the evidence accumulation process during decision making. Drift rates were larger when the processing of rule- and belief-based problem features converged, whereas they were smaller when they diverged. In addition, we found that only a limited number of individuals adjusted their decision criteria to process problems with a conflict between rule- and belief-based information more carefully.

By using innovative analytical approaches such as mathematical models of decision making, researchers are able to use the complete

---

[8] A list of specific items included in the two studies and their German translations will be provided upon request.

information of the entire response time distributions of correct and incorrect responses to infer latent process parameters describing decision making and conflict resolution. Hence, using diffusion models may help to disambiguate inconsistent results regarding the effect of increased processing demands and varying task instructions on response times in belief bias tasks (Handley et al., 2011; Thompson et al., 2011; Trippas et al., 2017).

In the second study, we found that when assessing problems based on the believability of their conclusions, participants did not only show higher drift rates (i.e., a greater speed and efficiency of information-processing), but also a shift in decision criteria that made them respond less cautiously under belief than under logic instruction. Because variations in drift rate and boundary separation parameters may affect different parts of the response time distribution (Ratcliff, Schmiedek, & McKoon, 2008), previous studies may have found conflicting results considering only effects on mean response times. This benefit of distinguishing between the speed of evidence accumulation and decision cautiousness is particularly useful for individual differences research, as individuals with higher abilities have been shown to strategically adjust their decision criteria differently from individuals with lower abilities to adapt to task demands (Draheim et al., 2016, 2019; Schmiedek et al., 2007). We found some evidence for a relationship between cognitive abilities and decision boundaries in Study 1 ($r = -0.22$), but not in Study 2 (all $rs \leq |0.05|$).

Other innovative modeling approaches that have been used to shed light on transitive reasoning performance include signal detection theory (Dube, Rotello, & Heit, 2010), equal-variance signal detection theory (Trippas et al., 2018), and multinomial models (Klauer, 2010; Klauer, Musch, & Naumer, 2000). Previous studies using variations of signal detection theory have called into question whether a conflict between rule- and belief-based problem features affects the ability to discriminate between valid and invalid syllogisms (Dube et al., 2010; Trippas et al., 2018). Those results are only partly in line with ours, as we found that the large majority of participants showed an effect of congruency on drift rates (i.e., discrimination ability). One reason for this divergence may be found in the simultaneous consideration of response times and accuracies in the present study, whereas signal detection models only consider accuracy rates irrespective of response speed. It would be interesting to see if analyses based on multinomial models that may be extended to fit response time distributions arrive at similar conclusions (Heck & Erdfelder, 2016; Klauer & Kellen, 2018). In addition, it would be interesting to compare our results with results from diffusion models that specifically model the conflict between rule- and belief-based processing. However, dual-process diffusion models have so far only been used as cognitive process models but not as measurement models, and can therefore only predict empirical data but not fit them (Alós-Ferrer, 2018; Caplin & Martin, 2016). In comparison, the diffusion model for conflict tasks has been developed to fit empirical data (Ulrich, Schröter, Leuthold, & Birngruber, 2015), but will likely need many more trials than used in the present study to describe the conflict resolution process reliably (White, Servant, & Logan, 2018).

### 6.1. Cognitive abilities and thinking styles affect transitive reasoning performance through different mechanisms

Our results strengthen the conceptual distinction between the algorithmic and reflective mind (Stanovich, 2009, 2010), as we found suggestive evidence for a double dissociation in the way that individual differences in algorithmic and reflective thinking moderated the processing of conflicting rule- and belief-based information. On the one hand, intelligent individuals seemed to quickly generate abstract representations of reasoning problems bereft of low-level intrinsic problem features. This impaired their ability to evaluate the believability of conclusions, either because they could not access these low-level intrinsic problem features as readily as individuals who did not automatically strip them from their mental problem representations or

because these quickly generated abstract representations interfered with simultaneously held belief-based representations. On the other hand, individuals with a disposition to engage in reflective thinking seemed to be as likely to generate a belief-based default response as other individuals, but inhibited and overrode the incorrect belief-based response more quickly and successfully. Hence, they showed no disadvantage when asked to assess problems based on their believability, but outperformed individuals with a lower need for cognition when asked to assess problems based on their logical validity.

Taken together, and to the extent that fluid intelligence and need for cognition are involved in the key functions of the algorithmic mind and the reflective mind (Stanovich, 2009, 2010), these results confirm the idea that algorithmic and reflective thinking abilities affect the processing of conflict information through different mechanisms and at different stages of information-processing. In particular, our results suggest that algorithmic abilities may facilitate the creation of decoupled problem representations (for similar findings, see Thompson et al., 2018; Thompson & Johnson, 2014), whereas a disposition to engage in critical thinking may facilitate the successful override of Type 1 responses. As such, our results are consistent with previous research that suggested that different individuals approach and solve reasoning tasks in qualitatively different ways (Erceg et al., 2019). All in all, results from individual-differences studies suggest that in different individuals, performance in reasoning tasks may be affected by different cognitive processes. Crucially, only individual differences in the reflective, but not in the algorithmic mind, predicted successful conflict resolution in the standard belief bias condition. Across both studies, we found that individual differences in algorithmic abilities were instead positively related to the general efficiency of information processing irrespective of congruence or incongruence between rule- and belief-based problem features.

### 6.2. Implications for dual-process accounts of reasoning

Our results contribute to the growing body of evidence challenging standard dual-process theories of reasoning (e.g., De Neys, 2012, 2014; Evans et al., 2009; Handley et al., 2011; Handley & Trippas, 2015; Newman et al., 2017; Pennycook et al., 2015; Thompson et al., 2018; Trippas et al., 2017). In particular, there are two findings that are not easily reconciled with the default interventionist account (Evans, 2008). First, standard dual-process theories cannot explain why participants asked to evaluate the believability of a conclusion performed worse when rule- and belief-based problem features were in conflict than when they were aligned. Second, they cannot explain why this impairment in performance was even greater for more intelligent individuals than less intelligent ones, as during initial stages of reasoning only belief-based default responses should be generated, which should not be harder to access for more intelligent reasoners. Instead, both findings are consistent with recent theoretical developments proposing that rule- and belief-based problem features are processed in parallel and that reasoning problems that elicit a conflict between rule- and belief-based may also elicit more than one Type 1 response (De Neys, 2012, 2014; Handley & Trippas, 2015; Pennycook et al., 2015). However, our results may also be reconciled with a modified default-interventionist model that assumes that different individuals generate different Type 1 default responses. As such, our results do not pose a severe challenge to the basic architecture and processing assumptions of the default interventionist account, but rather suggest that the default may vary between persons.

### 6.3. Limitations

One limitation of our studies is that we used ability measures for assessing individual differences in algorithmic thinking ability and self-report measures for assessing individual differences in reflective thinking. These measures are widely used to assess the two constructs in reasoning research, which allowed us to directly relate results from our study to a large number of previous findings. Nevertheless, it is important to note that constructs were confounded with measurement methods in the present studies. This is particularly problematic as ability and self-report measures of several constructs, including intelligence, self-control, self-regulation, prospective memory, and many more, have been repeatedly shown to be only weakly related (Eisenberg et al., 2019; Hedge, Powell, Bompas, & Sumner, 2020; Herreen & Zajac, 2017; Jacobs & Roodenburg, 2014; Pennycook, Ross, Koehler, & Fugelsang, 2017; Saunders, Milyavskaya, Etz, Randles, & Inzlicht, 2018; Uttl & Kibreab, 2011; Wennerhold & Friese, 2020). Future research should therefore strive to disambiguate this confounding factor and include objective measures of reflective thinking such as probabilistic reasoning and counter-example seeking.

Another limitation is the composition of our samples, which consisted mostly of university students. Range restriction in cognitive abilities could therefore have affected the size and direction of correlations between covariates and experimental effects. To prevent this, we recruited participants from a broad range of majors and included only a limited number of participants who were currently enrolled in highly competitive majors. The success of our recruitment strategy is reflected in mean performance in the APM, which, when extrapolated from 18 to 36 items, corresponds to an average of 22 correctly solved items. According to the norms presented in the manual, participants between 23 and 27 years who solved 22 items were placed in the 50th percentile of the ability distribution (Raven et al., 1994). This suggests that our sample was not particularly restricted in cognitive abilities.

A further limitation of the present study lies in the relative weak evidence we found for a full double dissociation between cognitive and reflective abilities. Whereas we could show that participants' need for cognition was more predictive of successfully ignoring *interfering belief-based* information than their intelligence, we could not conclusively demonstrate that participants' intelligence was more predictive of difficulties ignoring *interfering rule-based information* than their need for cognition. Many factors, including the low power of the *z*-test for comparing correlations, the measurement of cognitive and reflective abilities on a manifest variable level, and the moderate sample size, may have limited the conclusiveness of our results. We therefore suggest that future studies build upon our approach and measure cognitive and reflective abilities with multiple indicators in larger (and possibly more heterogeneous) samples, allowing the testing of latent correlations and latent correlation differences. Ideally, model parameters and their associations with cognitive and reflective abilities should then be estimated in a hierarchical Bayesian cognitive latent variable framework, as this would further increase the stability of latent parameter estimates (Schubert et al., 2019; Vandekerckhove, 2014).

### 6.4. Conclusions

By using the diffusion model, we simultaneously took into account both accuracy rates and response times to describe transitive reasoning performance. This enabled us to distinguish between experimental effects on the speed of evidence accumulation, on the speed of non-decisional processes, and on decision criteria. Across two studies we found that individual differences in need for cognition predicted successful conflict resolution under logic instruction, as reflected in the decrease in drift rates from non-conflict to conflict trials. This suggests that a disposition to engage in reflective thinking facilitates the inhibition and successful override of Type 1 responses. Intelligence, however, was negatively related to successful conflict resolution under belief instruction, which suggests that individuals with high cognitive abilities quickly generated a higher-level logical response that interfered with their ability to evaluate lower-level intrinsic problem features. Taken together, this tentative double dissociation suggests that algorithmic and reflective thinking abilities may affect the processing of conflict information through different mechanisms and at different stages of information-processing: Greater cognitive abilities facilitate the creation

of decoupled problem representations, whereas a greater disposition to engage in critical thinking facilitates the successful override of Type 1 responses.

## Acknowledgements

## Appendix A. Supplementary data

The data supporting the findings of the study are available in the Open Science Framework repository at https://osf.io/ycb2t/. Supplementary data to this article can be found online at [https://doi.org/10.1016/j.cognition.2021.104629].

## References

Akaike, H. (1973). In B. N. Petrov, & F. Csaki (Eds.), *Information theory and an extension of the maximum likelihood principle* (pp. 267–281). Akademiai Kiado.

Alós-Ferrer, C. (2018). A dual-process diffusion model. *Journal of Behavioral Decision Making, 31*(2), 203–218. https://doi.org/10.1002/bdm.1960.

Aschenbrenner, A. J., Balota, D. A., Gordon, B. A., Ratcliff, R., & Morris, J. C. (2016). A diffusion model analysis of episodic recognition in individuals with a family history for Alzheimer disease: The adult children study. *Neuropsychology, 30*(2), 225–238. https://doi.org/10.1037/neu0000222.

Ball, L. J., Thompson, V. A., & Stupple, E. J. N. (2018). Conflict and dual process theory: The case of belief bias. In *Dual process theory 2.0* (pp. 100–120). Routledge/Taylor & Francis Group.

Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology, 19*(2), 137–181. https://doi.org/10.1016/0022-2496(79)90016-6.

Banks, A. P., & Hope, C. (2014). Heuristic and analytic processes in reasoning: An event-related potential study of belief bias. *Psychophysiology, 51*(3), 290–297. https://doi.org/10.1111/psyp.12169.

Bialek, M., & Pennycook, G. (2018). The cognitive reflection test is robust to multiple exposures. *Behavior Research Methods, 50*(5), 1953–1959. https://doi.org/10.3758/s13428-017-0963-x.

Blacksmith, N., Yang, Y., Behrend, T. S., & Ruark, G. A. (2019). Assessing the validity of inferences from scores on the cognitive reflection test. *Journal of Behavioral Decision Making, 32*(5), 599–612. https://doi.org/10.1002/bdm.2133.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153–178. https://doi.org/10.1016/j.cogpsych.2007.12.002.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116.

Caplin, A., & Martin, D. (2016). The dual-process drift diffusion model: Evidence from response times. *Economic Inquiry, 54*(2), 1274–1282. https://doi.org/10.1111/ecin.12294.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press. https://doi.org/10.1017/CBO9780511571312.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41*(4), 244–262. https://doi.org/10.1016/j.intell.2013.04.003.

Chuderski, A. (2014). The relational integration task explains fluid reasoning above and beyond other working memory tasks. *Memory & Cognition, 42*(3), 448–463. https://doi.org/10.3758/s13421-013-0366-x.

Chuderski, A., & Jastrzębski, J. (2018). Much ado about aha!: Insight problem solving is strongly related to working memory capacity and reasoning ability. *Journal of Experimental Psychology. General, 147*(2), 257–281. https://doi.org/10.1037/xge0000378.

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology, 51*(2), 291–294. https://doi.org/10.1037/h0042761.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17*(5), 428–433. https://doi.org/10.1111/j.1467-9280.2006.01723.x.

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science, 7*(1), 28–38. https://doi.org/10.1177/1745691611429354.

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning, 20*(2), 169–187. https://doi.org/10.1080/13546783.2013.854725.

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case

example. *Perspectives on Psychological Science.* https://doi.org/10.1177/1745691616690990.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin, 145*(5), 508–535. https://doi.org/10.1037/bul0000192.

Dube, C., Rotello, C. M., & Heit, E. (2010). Assessing the belief bias effect with ROCs: It's a response bias effect. *Psychological Review, 117*(3), 831–863. https://doi.org/10.1037/a0019634.

Eisenberg, I. W., Bissett, P. G., Zeynep Enkavi, A., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications, 10*(1), 2319. https://doi.org/10.1038/s41467-019-10301-1.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, 71*(2), 390–405. https://doi.org/10.1037/0022-3514.71.2.390.

Erceg, N., Galic, Z., & Bubić, A. (2019). *Who detects and why? Individual differences in abilities, knowledge and thinking dispositions among different types of problem solvers and their implications for the validity of reasoning tasks* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/w5zau.

Erceg, N., Galic, Z., & Ružojčić, M. (2020). *A reflection on cognitive reflection – testing convergent validity of two versions of the cognitive reflection test.* https://doi.org/10.31234/osf.io/ewrtq.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629.

Evans, J. S. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11*(3), 295–306. https://doi.org/10.3758/BF03196976.

Evans, J. S. B. T., Handley, S. J., & Bacon, A. M. (2009). Reasoning under time pressure. *Experimental Psychology, 56*(2), 77–83. https://doi.org/10.1027/1618-3169.56.2.77.

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., & Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: A structural model-based approach. *Journal of Neuroscience, 31*(47), 17242–17249. https://doi.org/10.1523/JNEUROSCI.0309-11.2011.

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

Frischkorn, G. T., & Schubert, A.-L. (2018). Cognitive models in intelligence research: Advantages and recommendations for their application. *Journal of Intelligence, 6*(3), 34. https://doi.org/10.3390/jintelligence6030034.

Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence, 52*, 71–79. https://doi.org/10.1016/j.intell.2015.07.006.

Handley, S. J., & Trippas, D. (2015). Dual processes and the interplay between knowledge and structure: A new parallel processing model. *Psychology of Learning and Motivation - Advances in Research and Theory, 62*, 33–58. https://doi.org/10.1016/bs.plm.2014.09.002.

Handley, S. J., Newstead, S. E., & Trippas, D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(1), 28–43. https://doi.org/10.1037/a0021098.

Heck, D. W., & Erdfelder, E. (2016). Extending multinomial processing tree models to measure the relative speed of cognitive processes. *Psychonomic Bulletin & Review, 23*(5), 1440–1465. https://doi.org/10.3758/s13423-016-1025-6.

Hedge, C., Powell, G., Bompas, A., & Sumner, P. (2020). Self-reported impulsivity does not predict response caution. *Personality and Individual Differences, 167*, 110257. https://doi.org/10.1016/j.paid.2020.110257.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*. https://doi.org/10.3389/fnins.2014.00150.

Herreen, D., & Zajac, I. T. (2017). The reliability and validity of a self-report measure of cognitive abilities in older adults: More personality than cognitive function. *Journal of Intelligence, 6*(1). https://doi.org/10.3390/jintelligence6010001.

Howarth, S., Handley, S. J., & Walsh, C. (2016). The logic-bias effect: The role of effortful processing in the resolution of belief-logic conflict. *Memory & Cognition, 44*(2), 330–349. https://doi.org/10.3758/s13421-015-0555-x.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*(3), 702–721. https://doi.org/10.3758/s13428-013-0411-5.

Jacobs, K. E., & Roodenburg, J. (2014). The development and validation of the self-report measure of cognitive abilities: A multitrait–multimethod study. *Intelligence, 42*, 5–21. https://doi.org/10.1016/j.intell.2013.09.004.

Keller, J., Bohner, G., & Erb, H.-P. (2000). Intuitive und heuristische Urteilsbildung—Verschiedene Prozesse? Präsentation einer deutschen Fassung des "rational–experiential inventory" sowie neuer Selbstberichtskalen zur Heuristiknutzung. [intuitive und heuristic judgment—Different processes? Presentation of a German version of the rational–experiential inventory and of new self-report scales of heuristic use.]. *Zeitschrift für Sozialpsychologie, 31*(2), 87–101. https://doi.org/10.1024//0044-3514.31.2.87.

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75*(1), 70–98. https://doi.org/10.1007/s11336-009-9141-0.

Klauer, K. C., & Kellen, D. (2018). RT-MPTs: Process models for response-time distributions based on multinomial processing trees with applications to recognition

memory. *Journal of Mathematical Psychology, 82*, 111–130. https://doi.org/10.1016/j.jmp.2017.12.003.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*(4), 852–884. https://doi.org/10.1037/0033-295X.107.4.852.

von Krause, M., Lerche, V., Schubert, A.-L., & Voss, A. (2020). Do non-decision times mediate the association between age and intelligence across different content and process domains? *Journal of Intelligence, 8*(3), 33. https://doi.org/10.3390/jintelligence8030033.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*(4), 389–433. https://doi.org/10.1016/S0160-2896(05)80012-1.

Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01324.

Lerche, V., & Voss, A. (2017). Retest reliability of the parameters of the Ratcliff diffusion model. *Psychological Research, 81*(3), 629–652. https://doi.org/10.1007/s00426-016-0770-5.

Lerche, V., & Voss, A. (2019). Experimental validation of the diffusion model based on a slow response time paradigm. *Psychological Research, 83*(6), 1194–1209. https://doi.org/10.1007/s00426-017-0945-8.

Lerche, V., von Krause, M., Voss, A., Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2020). Diffusion modeling and intelligence: Drift rates show both domain-general and domain-specific relations with intelligence. *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0000774.

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2016). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*. https://doi.org/10.1177/0272989X0102100105.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6*(3), 312–319. https://doi.org/10.3758/BF03197461.

van der Maas, H. L. J., Molenaar, D., Maris, G., Kievit, R. A., & Borsboom, D. (2011). Cognitive psychology meets psychometric theory: On the relation between process models for decision making and latent variable models for individual differences. *Psychological Review, 118*(2), 339–356. https://doi.org/10.1037/a0022749.

Newman, I. R., Gibb, M., & Thompson, V. A. (2017). Rule-based reasoning is fast and belief-based reasoning can be slow: Challenging current explanations of belief-bias and base-rate neglect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(7), 1154–1170. https://doi.org/10.1037/xlm0000372.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence, 36*(6), 641–652. https://doi.org/10.1016/j.intell.2008.01.007.

Pennycook, G., & Ross, R. M. (2016). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00009.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34–72. https://doi.org/10.1016/j.cogpsych.2015.05.001.

Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2017). Dunning-Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin & Review, 24*(6), 1774–1784. https://doi.org/10.3758/s13423-017-1242-7.

Pew, R. W. (1969). The speed-accuracy operating characteristic. *Acta Psychologica, 30*, 16–26. https://doi.org/10.1016/0001-6918(69)90035-3.

Piaget, J. (1928). *Judgment and reasoning in the child* (pp. viii, 260). Brace: Harcourt. https://doi.org/10.4324/9780203207260.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry* (pp. vii, 411). Basic Books.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review, 9*(2), 278–291. https://doi.org/10.3758/BF03196283.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420.

Ratcliff, R., Schmiedek, F., & McKoon, G. (2008). A diffusion model explanation of the worst performance rule for reaction time and IQ. *Intelligence, 36*(1), 10–17. https://doi.org/10.1016/j.intell.2006.12.002.

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology, 60*(3), 127–157. https://doi.org/10.1016/j.cogpsych.2009.09.001.

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General, 140*(3), 464–487. https://doi.org/10.1037/a0023810.

Raven, J. C., Court, J. H., & Raven, J. (1994). *Manual for Raven's progressive matrices and mill hill vocabulary scales. Advanced progressive matrices*. Oxford University Press.

van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review, 24*(2), 547–556. https://doi.org/10.3758/s13423-016-1081-y.

Saunders, B., Milyavskaya, M., Etz, A., Randles, D., & Inzlicht, M. (2018). Reported self-control is not meaningfully associated with inhibition-related executive function: A

Bayesian analysis. *Collabra: Psychology, 4*(1), 39. https://doi.org/10.1525/collabra.134.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136*(3), 414–429. https://doi.org/10.1037/0096-3445.136.3.414.

Schmitz, F., & Wilhelm, O. (2016). Modeling mental speed: Decomposing response time distributions in elementary cognitive tasks and correlations with working memory capacity and fluid intelligence. *Journal of Intelligence, 4*(4), 13. https://doi.org/10.3390/jintelligence4040013.

Schubert, A.-L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence, 51*, 28–46. https://doi.org/10.1016/j.intell.2015.05.002.

Schubert, A.-L., Frischkorn, G. T., Hagemann, D., & Voss, A. (2016). Trait characteristics of diffusion model parameters. *Journal of Intelligence, 4*(3), 7. https://doi.org/10.3390/jintelligence4030007.

Schubert, A.-L., Nunez, M. D., Hagemann, D., & Vandekerckhove, J. (2019). Individual differences in cortical processing speed predict cognitive abilities: A model-based cognitive neuroscience account. *Computational Brain & Behavior, 2*(2), 64–84. https://doi.org/10.1007/s42113-018-0021-5.

Schubert, A.-L., Hagemann, D., Löffler, C., & Frischkorn, G. T. (2020). Disentangling the effects of processing speed on the association between age differences and fluid intelligence. *Journal of Intelligence, 8*(1), 1. https://doi.org/10.3390/jintelligence8010001.

Schubert, A.-L., Ferreira, M. B., Mata, A., & Riemenschneider, B. (2021, February 2). *A diffusion model analysis of belief bias: Different cognitive mechanisms explain how cognitive abilities and thinking styles contribute to conflict resolution in reasoning.* 10.17605/OSF.IO/YCB2T.

Schulz-Zhecheva, Y., Voelkle, M. C., Beauducel, A., Biscaldi, M., & Klein, C. (2016). Predicting fluid intelligence by components of reaction time distributions from simple choice reaction time tasks. *Journal of Intelligence, 4*(3), 8. https://doi.org/10.3390/jintelligence4030008.

Shipley, W. C., Christian, P. G., Martin, T. A., & Klein, A. M. (2009). *Shipley-2. Western psychological services*.

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In K. Frankish, & J. S. B. T. Evans (Eds.), *Two minds: Dual processes and beyond* (pp. 55–88). Oxford University Press.

Stanovich, K. E. (2010). *Rationality and the reflective mind.* Oxford University Press.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645–665. https://doi.org/10.1017/S0140525X00003435.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The rationality quotient: Toward a test of rational thinking* (1st ed.). MIT Press Ltd.

Sternberg, R. J. (1980a). The development of linear syllogistic reasoning. *Journal of Experimental Child Psychology, 29*(2), 340–356. https://doi.org/10.1016/0022-0965(80)90025-9.

Sternberg, R. J. (1980b). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109*(2), 119–159.

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning, 20*(2), 215–244. https://doi.org/10.1080/13546783.2013.869763.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology, 63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001.

Thompson, V. A., Pennycook, G., Trippas, D., & Evans, J. S. B. T. (2018). Do smart people have better intuitions? *Journal of Experimental Psychology: General, 147*(7), 945–961. https://doi.org/10.1037/xge0000457.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275. https://doi.org/10.3758/s13421-011-0104-1.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning, 20*(2), 147–168. https://doi.org/10.1080/13546783.2013.844729.

Trippas, D., Thompson, V. A., & Handley, S. J. (2017). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory & Cognition, 45*(4), 539–552. https://doi.org/10.3758/s13421-016-0680-1.

Trippas, D., Kellen, D., Singmann, H., Pennycook, G., Koehler, D. J., Fugelsang, J. A., & Dubé, C. (2018). Characterizing belief bias in syllogistic reasoning: A hierarchical Bayesian meta-analysis of ROC data. *Psychonomic Bulletin & Review, 25*(6), 2141–2174. https://doi.org/10.3758/s13423-018-1460-7.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5.

Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2015). Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology, 78*, 148–174. https://doi.org/10.1016/j.cogpsych.2015.02.005.

Uttl, B., & Kibreab, M. (2011). Self-report measures of prospective memory are reliable but not valid. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale, 65*(1), 57–68. https://doi.org/10.1037/a0022843.

Vandekerckhove, J. (2014). A cognitive latent variable model for the simultaneous analysis of behavioral and personality data. *Journal of Mathematical Psychology, 60*, 58–71. https://doi.org/10.1016/j.jmp.2014.06.004.

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39*(4), 767–775. https://doi.org/10.3758/BF03192967.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology, 60*(6), 385–402. https://doi.org/10.1027/1618-3169/a000218.

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00336.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11*(1), 192–196. https://doi.org/10.3758/BF03206482.

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition, 40*(2), 145–160. https://doi.org/10.3758/s13421-011-0158-0.

Wennerhold, L., & Friese, M. (2020). Why self-report measures of self-control and inhibition tasks do not substantially correlate. *Collabra: Psychology, 6*(1), 9. https://doi.org/10.1525/collabra.276.

White, C. N., Servant, M., & Logan, G. D. (2018). Testing the validity of conflict drift-diffusion models for use in estimating cognitive processes: A parameter-recovery study. *Psychonomic Bulletin & Review, 25*(1), 286–301. https://doi.org/10.3758/s13423-017-1271-2.

Wilhelm, O., Hildebrandt, A. H., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00433.

Wright, B. C. (2001). Reconceptualizing the transitive inference ability: A framework for existing and future research. *Developmental Review, 21*(4), 375–422. https://doi.org/10.1006/drev.2000.0525.

Wright, B. C., & Smailes, J. (2015). Factors and processes in children's transitive deductions. *Journal of Cognitive Psychology, 27*(8), 967–978. https://doi.org/10.1080/20445911.2015.1063641.