COGNITIVE NEUROPSYCHOLOGY
2021, VOL. 38, NO. 6, 365–386
https://doi.org/10.1080/02643294.2022.2041589

Routledge
Taylor & Francis Group

Can a fast thinker be a good thinker? The neural correlates of base-rate neglect measured using a two-response paradigm

Oshin Vartanian [a,b], Timothy K. Lam[a], Elaine Maceda[a] and Wim De Neys[c]

[a]Human Effectiveness Section, Defence Research and Development Canada, Toronto, ON, Canada; [b]Department of Psychology, University of Toronto, Toronto, ON, Canada; [c]Laboratory for the Psychology of Child Development and Education, Centre National de la Recherche Scientifique, Université de Paris, Paris, France

ABSTRACT

Traditionally, it has been assumed that logical thinking requires deliberation. However, people can also make logical responses quickly, exhibiting *logical intuitions*. We examined the neural correlates of logical intuitions by administering base rate problems during fMRI scanning using a two-response paradigm where participants first responded quickly and then reflectively to problems that did or did not pit a normative response against an intuitively-cued stereotypical response (i.e., conflict vs. non-conflict problems). As predicted, participants were less likely to make judgments in accordance with base rates on conflict problems. Critically, in only 4% of cases did longer deliberation change an initially biased response to a normatively correct response. The fMRI data revealed that intuitively-made initial biased judgments nevertheless activate regions typically involved in cognitive control, executive functions and attention, including anterior, inferior, middle and superior frontal cortex, suggesting that even when errors are made, there might be very early awareness of conflict.

KEYWORDS

Heuristics and biases; base-rate neglect; two-response paradigm

1. Introduction

One of the most enduring beliefs in psychology involves the idea that logical thinking requires effort, consistent with the common belief that following logical or mathematical rules is hard (Kahneman, 2011). This idea is perhaps most evident in dual-process theories of thinking, reasoning, and judgment and decision making where an intuitive mode of thinking (System 1) is contrasted with a deliberative mode of thinking (System 2). Although dual-process theories come in many different flavours, a central theme that runs through this literature is that thinking is susceptible to bias in cases where a conflict exists between an intuitively-cued but logically incorrect response, and a normatively correct response that requires deliberation for implementation. Consider the following logical argument (i.e., syllogism) that consists of two premises followed by a conclusion: All things that are smoked are good for the health; Cigarettes are smoked; Conclusion: Cigarettes are good for the health (taken from Toplak et al., 2014). Although by virtue of its format this is a valid argument, people tend not to accept the conclusion as valid because its content contradicts with their beliefs. In such cases, unless motivation and cognitive ability exist in support of a reflective thinking style to override the intuitively-cued response, one's thinking will be led astray, leading to biased (i.e., logically and/or normatively incorrect) responding.

Indeed, many standard paradigms in experimental psychology (e.g., belief bias, base-rate neglect, ball-and-bat problems) have been designed specifically to instantiate a conflict between intuitive and deliberative modes of thinking (see De Neys & Pennycook, 2019). The results from this large body of literature have demonstrated that in many cases thinking in accordance with logical and probabilistic principles requires time-consuming, deliberative processing (e.g., Evans, 2008; Stanovich & West, 2000). Furthermore, because there is a tendency to minimize demanding computations, most participants tend to opt for the intuitively-cued responses, rather than engage in the additional information processing necessary for logical thought (Evans & Stanovich, 2013; Kahneman, 2011). As such, unless such

CONTACT Oshin Vartanian oshin.vartanian@drdc-rddc.gc.ca Human Effectiveness Section, Defence Research and Development Canada, Toronto, ON, Canada; Toronto Research Centre, Toronto, M3 K 2C9 Canada

deliberate computations are performed to override the intuitively-cued response, people will not respond in accordance with logical and probabilistic norms (Stanovich & West, 2000).

An important feature of such dual-process accounts is that logical and normative thinking requires time. Guo et al. (2017) made an important contribution to this literature by focusing on the impact of time pressure on a classic effect from judgment and decision making, namely framing. Specifically, in accordance with theories of rational choice, a person's choices should be *description invariant*—meaning that if the expected value of two outcomes is the same, then one's preference for those outcomes should not vary depending on the description of the problem. However, there is substantial evidence to suggest that in the context of risky choices, people's preferences are affected by the description of the problem (Kahneman & Tversky, 1979; Tversky & Kahneman, 1981). In the context of three experiments, Guo et al. (2017) used time pressure to assess the accuracy of two competing explanations of this framing effect. On the one hand, some have argued that the framing effect arises as a function of a deliberative process, growing larger with time. On the other hand, in accordance with dual-process theory, framing effects are hypothesized to occur due to an intuitive mode of thought that responds automatically to stimuli. By systematically manipulating time pressure, the researchers were able to show that framing effects increased under time pressure, supporting the dual-process account. Their results demonstrated that people's thinking is more likely to deviate from the axioms of rational thinking if they are given less time for deliberation (see also Evans & Curtis-Holmes, 2005; Lawson et al., 2020).

1.1. Awareness of conflict

Despite the evidence presented above, some researchers have begun to reconsider some key tenets of the dual-process account. First, it is generally assumed that in heuristics and biases tasks, biased responding occurs outside of the window of awareness. In other words, people violate logical and probabilistic norms without being aware that they have done so. However, there is now growing evidence to suggest that this might not necessarily be the case. For example, behavioural studies have demonstrated that participants have lower confidence in their choices and deliberate longer when they make normative errors in conflict problems (De Neys, 2012; De Neys & Glumicic, 2008). Such findings demonstrate that biased reasoners exhibit some sensitivity to rule violations. Importantly, this sensitivity to conflict is also apparent under time pressure and cognitive load (De Neys, 2017), suggesting that even when deliberative processing is not possible, sensitivity to conflict is nevertheless present.

Supporting these behavioural results, there is now also evidence from brain imaging studies to suggest that biased reasoners might be aware of normative rule violations. De Neys et al. (2008) presented participants with various versions of base rate problems in the magnetic resonance imaging (MRI) scanner while functional scans were collected. On critical conflict problems that pitted a normative response against an intuitively-cued biased response, they found that the anterior cingulate cortex (ACC)—a region strongly linked to error detection and monitoring (e.g., Botvinick et al., 2004; Ridderinkhof et al., 2004a; van Veen & Carter, 2006)—was activated even on trials when participants made normatively incorrect responses. From a neurological perspective, this suggests that biased reasoners exhibited sensitivity to the presence of conflict, despite their inability to override it. In contrast, on trials where normative responses were registered, in addition to the ACC, the right inferior frontal gyrus was also activated, a region strongly linked with cognitive, behavioural, and emotional inhibition (Aron et al., 2003, 2004, 2014). This suggests that responding in accordance with statistical norms requires not only sensitivity to conflict, but in addition, the ability to override the intuitively-cued biased response in favour of rational norms.

There are two overarching conclusions that can be drawn from this body of work. First, in cases where there is a conflict between a normative response and an intuitively-cued biased response, responding in accordance with logical norms necessitates two abilities: The ability to detect the presence of conflict, as well as the ability to override the intuitively-cued biased response in favour of the normative response. The neurological evidence in support of this argument is provided by the presence of activation in the ACC as well as in the inferior frontal gyrus on trials when participants make the normative response on conflict trials (De Neys et al., 2008).

Specifically, this activation pattern is consistent with the role of the ACC in conflict detection and error monitoring, as well as the inferior frontal gyrus' role in inhibition. Thus, this body of work suggests that biased responding does not necessarily result from a failure to detect conflict. Rather, participants might lack the necessary tools to override the conflict in favour of normative responding. Second, and of particular relevance for the present purposes, the ACC was activated even on trials when participants made normatively *incorrect* responses—defined as responses that deviated from statistical norms (i.e., base rates). What this finding suggests is that the inability to respond normatively does not necessarily mean that participants were unaware of the presence of conflict. As such, the ACC activation observed on such trials is inconsistent with theories that posit that biased responding necessarily occurs outside of the window of awareness.

1.2. Logical intuitions

A second line of evidence that has also led to a reconsideration of the dual-process account presented above is that people appear to have the ability to make sound logical or normative choices *intuitively* (i.e., quickly and without deliberation). Empirical support for such "logical intuitions" has come mainly from the two-response paradigm, which was designed to explore the time course of intuitive and deliberative processing (Thompson et al., 2011). In this paradigm, and in response to problems that pit a normative response against an intuitively-cued biased response, participants are given two opportunities to respond—once quickly based on their intuitions, and then again following further deliberation. To ensure that the first response is generated intuitively, the experimenters make sure that it occurs under time pressure, under cognitive load, or both (Bago & De Neys, 2017; Newman et al., 2017). Based on the value assigned to deliberation for arriving at logical and/or normative choices, one would expect that participants would be more likely to arrive at such choices when they are given more rather than less time to deliberate on the problem. Similarly, one might expect to see many instances in which upon further reflection, participants change an initially incorrect response to a logically or normatively correct response. In contrast to these predictions derived from dual-process theories, what the results of the two-response paradigm have shown is that participants who gave a logical/normative final response (i.e., following deliberation) had already done so in the initial response stage (Bago & De Neys, 2017, 2019a). This suggests that good reasoners do not necessarily deliberate better; in fact, it might be that they have better (logical) intuitions (Bago & De Neys, 2019a; Thompson et al., 2018).

Recent electrophysiological evidence derived from event-related potentials (ERP) has supported the notion of logical intuitions (Bago et al., 2018; Banks & Hope, 2014). Specifically, while recording their electroencephalogram (EEG), Bago et al. (2018) presented participants with base rate problems in which an intuitively-cued stereotypical response was either congruent (i.e., non-conflict problems) or incongruent (i.e., conflict problems) with the normative response that was cued by base rates. Because of the high temporal resolution of EEG, it is possible to determine with a high degree of precision when a conflict between base rates and intuitively-cued stereotypical responses has been detected. The results demonstrated that when base rates and stereotypical descriptions cued conflicting responses, there was increased centro-parietal N2 (175-250 ms time window) and frontal P3 (300-500 ms time window) activity. Critically, the increased N2 activity for conflict problems was observed regardless of whether the participants responded correctly in accordance with base rates or incorrectly according to intuitively-cued stereotypical information. The results from Bago et al. (2018) suggest that conflict sensitivity can occur very early, without necessarily the involvement of slow, deliberative processes. Importantly, conflict sensitivity is present even when participants have made normatively incorrect responses, suggesting that despite suboptimal performance they might nevertheless be aware of the presence of conflict.

1.3. Individual differences

An important and relatively recent development in the literature on heuristics and biases involves a growing appreciation of individual differences in susceptibility to biased reasoning. Specifically, a large body of literature has now shown that individuals with greater levels of cognitive capacity and/or ability and those with specific thinking styles are less likely to fall prey to errors in thinking in such tasks (Pennycook et al., 2014; Stanovich & West, 1998, 2000, 2008; Toplak et al., 2011). For example, Toplak et al. (2011) examined the contributions of cognitive ability and thinking style to performance on a wide host of tasks involving probabilistic reasoning, hypothetical thought, theory justification, scientific reasoning, and the tendency to think statistically. Cognitive ability was measured using the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999), whereas the disposition to think analytically was measured using the Cognitive Reflection Test (CRT; Frederick, 2005)—an instrument designed to measure the propensity to suppress a fast intuitive response in favour of a reflective, deliberative response. The host of tasks that the experimenters focused on were selected because they cued an intuitive but incorrect response that needed to be overcome for normative/logical responding. They found that CRT scores were a unique predictor of performance on such heuristics and biases tasks after controlling for cognitive ability. Pennycook et al. (2014) reinforced this conclusion by demonstrating that both cognitive ability and CRT contribute positively to performance on base rate problems.

Recently, Vartanian et al. (2018) investigated the neural correlates of individual differences in performance on base rate problems. In Experiment 1, conducted outside of the MRI scanner, they demonstrated that CRT scores predicted performance on conflict problems, whereas short-term memory (STM) span predicted performance on non-conflict problems. Experiment 2, conducted in the MRI scanner, replicated this behavioural dissociation. Furthermore, the results demonstrated that conflict problems were associated with greater activation in the ACC—a key region for conflict detection—even in cases when participants responded stereotypically. This result replicated their earlier findings by demonstrating the presence of conflict sensitivity within ACC even in cases of stereotypical responding (see De Neys et al., 2008). In addition, from an individual-differences perspective, in participants with higher CRT scores, conflict problems were associated with greater activation in the posterior cingulate cortex (PCC), and activation in PCC covaried in relation to CRT scores during conflict problems. CRT scores also predicted activation in PCC in conflict problems (over and above non-conflict problems). Their results suggested that individual differences in reflective thinking are related to brain activation in PCC—a region involved in regulating attention between external and internal foci (Leech & Sharp, 2014).

The overall take-home message from this body of work is that because not everyone is susceptible to base-rate neglect to the same extent, individual differences need to be taken into account for a more complete understanding of the psychological and neurological underpinnings of this phenomenon. As emphasized by De Neys and Pennycook (2019), the examination of individual differences is also relevant to our understanding of conflict sensitivity. For example, although biased reasoners typically show some form of conflict detection, some individuals do not (Frey et al., 2018; Pennycook et al., 2015). One possibility is that such individuals might not have internalized or automatized the application of the necessary logical knowledge (Stanovich, 2018). Indeed, De Neys and Pennycook (2019) have argued that people's intuitive logical knowledge emerges from a learning process in which over time, key principles have been practiced to a point where they become automatic (De Neys, 2012; see also Kahneman, 2011; Shiffrin & Schneider, 1977; Stanovich, 2018). When this occurs, as is the case with good reasoners, the exercise of logical thinking becomes less effortful and can occur intuitively. Furthermore, because logical thinking in adulthood is correlated with a number of related abilities such as intelligence and executive functions (Stanovich & West, 1998, 2000, 2008; Toplak et al., 2011), it is likely that their development might benefit from similar conditions and exhibit similar trajectories. For example, it has been shown that people who exhibit higher intelligence as young adults exhibit a specific pattern of cortical maturation, characterized by a negative correlation between intelligence and cortical thickness in early childhood but a positive correlation in late childhood and beyond—particularly in the frontal lobes (Shaw et al., 2006). Because such longitudinal neurological data are currently not available in relation to reasoning, the developmental emergence of intuitive logical knowledge remains unknown. As such, further work is necessary to understand why some people can exhibit intuitive logic whereas others cannot. Here lesion studies could prove useful, as they can shed light on neural structures that might be necessary for the developmental emergence of logical competence, including intuitive logic (see Goel, 2007).

1.4. Present experiment

The aim of the present experiment was to examine the neural correlates of logical intuitions. Specifically, our experiment was designed to test two predictions derived from the literature discussed above. First, dual-process theories posit that thinking in accordance with rational norms requires effortful processing, which in turn necessitates time to implement. We will test that idea by examining whether people can perform in accordance with rational norms under time pressure, which limits the opportunity for effortful processing. Second, it has generally been assumed that biased responding occurs outside of the window of awareness. We will test that idea by examining whether activation in regions of the brain that are sensitive to the detection of error, cognitive control, executive functions and attention is nevertheless present when people register intuitively-cued biased responses, suggesting that there could be error sensitivity under such circumstances.

To do so, we administered base rate problems in the MRI scanner using a two-response paradigm. Participants were instructed to respond—first quickly within 3 s and then reflectively following 6 more seconds—to conflict and non-conflict problems. Conflict problems pit a normative response against an intuitively-cued stereotypical response, whereas non-conflict problems do not. In addition, we also administered a battery of cognitive ability (intelligence, simple working memory span [i.e., STM]), personality (Big Five), and thinking style (CRT) measures to explore their correlations with performance on conflict and non-conflict problems. We tested three hypotheses: First, behaviourally, we hypothesized that participants would exhibit logical intuitions. Logical intuitions would be reflected by similar rates of correct, normative responses regardless of time pressure, and a low rate of response switches following deliberation. Second, behaviourally, we hypothesized that superior performance on conflict problems, under time pressure or otherwise, would be correlated with individual differences in cognitive abilities and thinking styles. Third, from a neurological perspective, we hypothesized that participants would already exhibit conflict sensitivity when they are instructed to intuit and respond quickly, reflected by brain activation in regions that underlie error detection, attention and cognitive control on initial responses (compared to final responses). These regions have previously been shown to be activated during reasoning when participants' reflection is not restricted (e.g., De Martino et al., 2006; De Neys et al., 2008; Goel et al., 2000; Goel & Dolan, 2003; Houdé et al., 2011; Prado et al., 2011; Prado & Noveck, 2007; Stollstorff et al., 2012; Tsujii et al., 2010; Tsujii & Watanabe, 2010; Vartanian et al., 2018). If the activation is already observed for quick, initial responses, then this would indicate that neural regions involved in detecting and overcoming conflict can be engaged in short order and under time pressure (i.e., intuitively). Note that we did not predict that the aforementioned regions would be activated *exclusively* for the initial response and not for the final response. Rather, we predicted that their activation in that early phase would be consistent with the idea that sensitivity to conflict can be present early. Following Bago et al. (2018), we expected this intuitive or initial response conflict sensitivity to be present even when participants would generate a biased response on conflict problems. In other words, even in cases when people respond non-normatively, there may nevertheless be intuitive sensitivity to conflict.

2. Method

2.1. Participants

The protocol for this study was approved by the Human Research Ethics Committee of Defence Research and Development Canada (DRDC). Twenty-four neurologically healthy right-handed participants (21 males, 3 females) volunteered to participate in this study following informed consent ($M = 31.21$ years old, $SD = 10.08$).

2.2. Materials and procedure

Our experiment was conducted in two sessions. The first session involved the administration of all individual-differences measures to each participant in a single session in our laboratory at DRDC (Toronto Research Centre). We tested participants individually. We administered two measures of simple working memory span (i.e., STM tasks), modeled after Harrison et al.'s (2013) simple working memory span tasks. For word (verbal) span, four-letter monosyllabic words were presented one at a time on a monitor. After each block of words, participants were prompted by the software to recall the words they saw in the order in which they were presented. Blocks ranged from three to nine words. For spatial (matrix) span, participants were presented with a 4 × 4 matrix where one square (out of 16) appeared in red and the rest in white. At the end of each block of matrices, participants were instructed to recall the locations of the red squares in the order in which they were presented. Blocks ranged from three to nine matrices. The computer application provided a detailed description of each task. Before beginning the first trial, the experimenter reviewed the instructions and provided an example in each case to the participants. Note that both the verbal and matrix span are so-called simple working memory span tasks that primarily tax short-term memory storage capacity (e.g., Harrison et al., 2013; Cowan, 2008; Unsworth & Engle, 2007). In addition, to assess the inhibition component of executive functions, we administered the verbal version of the Stroop task (Stroop, 1935) using a computer software for administering cognitive tasks (Grushcow, 2008). For Stroop, the key metric was the difference in RT between the correct identification of incongruent trials (e.g., the word RED presented in blue) vs. the correct identification of congruent trials (e.g., the word RED presented in red).

Our measures of crystallized and fluid intelligence consisted of the Vocabulary (10 min) and Block Patterns (10 min) subsets of the Shipley-2, which were in turn standardized into full-scale intelligence scores (Shipley et al., 2009). We also administered the seven-item version of the CRT, which built on Frederick's (2005) original three-item version by adding four more items (Toplak et al., 2014). This instrument is designed to measure the propensity to suppress a fast, intuitive response in favour of a reflective, deliberative response (e.g., A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost? Incorrect answer = 10 cents; correct answer = 5 cents). CRT problems are believed to be ideal for probing the interplay between heuristic and analytic thinking precisely because the incorrect but intuitive response typically appears immediately, which in turn must be suppressed in favour of the correct and analytically-derived response.

Finally, the participants completed the Big Five Aspect Scales (BFAS, DeYoung et al., 2007). Although we administered the BFAS in its entirety, from a theoretical perspective we were only interested in the Conscientiousness factor that is derived from two aspect scales (i.e., Industriousness and Orderliness). Specifically, we reasoned that participants who score higher on Conscientiousness and are therefore more organized, diligent, and industrious might be more likely to invest more effort in deliberation when given the opportunity to reconsider an initial response.

Approximately 1–2 weeks after the first session, participants returned for their second session which involved the acquisition of functional MRI scans at York University's MRI Facility (https://mri.info.yorku.ca). The 48 base rate problems (24 conflict, 24 non-conflict) were selected from Pennycook et al.'s (2014) item pool, and were the same items administered by Vartanian et al. (2018). The following instructions were read verbatim by the experimenter as the participant viewed the text, prior to entering the scanner:

> In a big research project a large number of studies were carried out where a psychologist made short personality descriptions of the participants. In every study there were participants from two population groups (e.g., carpenters and policemen). In each study one participant was drawn at random from the sample. You'll get to see one personality trait of this randomly chosen participant. You'll also get information about the composition of the population groups tested in the study in question. You'll be asked to indicate to which population group the participant most likely belongs. You will be presented with 48 trials in the fMRI scanner, each representing a separate study.
>
> In the first stage (indicated by a green border around the question slide), we want you to respond with the very first answer that comes to mind. You don't need to think about it much. Just give the first answer that intuitively comes to mind as quickly as possible. You will have only 3 s to provide an answer!
>
> Then, in the second stage (indicated by a red border around the question slide), you will get more time (i.e., another 6 s) to reflect on the problem and enter your final response. In general this should allow you to make up your mind. There's no need to rush as much in the second stage. Once you have made up your mind after further deliberation, you can enter your response.

Your experimenter will review the instructions and answer any questions you might have before you start.

After confirming the participant understood the requirements of the task, two practice problems were presented for familiarization purposes. The following depicts a representative item from the 48-item set (presented in the fMRI scanner):

This study contains:
Lawyers and clowns
Person "L" is argumentative
There are 3 lawyers/997 clowns
Person "L" is more likely to be:
1) Clown
2) Lawyer

On all problems, the base rate contrast between the two categories was similarly extreme. After the participants exited the scanner, they were instructed to indicate their global confidence level in the accuracy of their judgments for initial and final responses using a 0–100 scale (0 = *not confident at all*, 100 = *extremely confident*).

Note that all problem content that was used in the present study was extensively pretested in advance (see Pennycook et al., 2015). Pennycook et al. made sure that the words that were selected to cue a stereotypical association did so consistently while avoiding extremely diagnostic cues. Such a non-extreme, moderate association is important. For convenience and consistency with prior work we label the response that is in line with the base rates as the correct, normative response. Critics of the base rate task (e.g., Gigerenzer et al., 1988) have long pointed out that if reasoners were to adopt a Bayesian approach and combine the base rate probabilities with the stereotypical description, then this can lead to interpretational complications when the description is extremely diagnostic. For example, imagine that we have an item with males and females as the two groups and give the description that Person "A" is "pregnant". Now, in this case, one would always need to conclude that Person "A" is a woman, regardless of the base rates. The more moderate descriptions in the present study help to avoid this potential problem. In addition, the extreme base rates (997/3) that were used in the current study further help to guarantee that even a very approximate Bayesian reasoner would need to pick the response cued by the base rates (see Bago & De Neys, 2017). Having said this, with more moderate base rates it is possible that incorrect responders will no longer show conflict sensitivity (e.g., Bago & De Neys, 2019b; Pennycook et al., 2015). In this sense moderate base rates tend to make the task more difficult, possibly also resulting in more conflict/control activation for correct responders.

2.3. Image acquisition and processing

Magnetic resonance images were acquired on a Siemens MAGNETOM Prisma Fit 3 Tesla system (Erlangen, Germany). We obtained T1-weighted anatomical images with the following parameters: repetition time = 2300 msec, echo time = 2.62 msec, and voxel size = $1 \times 1 \times 1$ mm^3, for a total of 192 axial slices covering the whole brain. For functional imaging, T2*-weighted gradient-echo images were acquired with the following parameters: repetition time = 1630 msec, echo time = 30 msec, flip angle = 66°, field of view = 240×240 mm^2, matrix = 120×120 voxels, and voxel size = $2 \times 2 \times 2$ mm^3, for a total of 72 contiguous 2-mm thick axial slices positioned to cover the whole brain. The first six volumes were removed to account for T1 equilibration effects. In total, 510 volumes were acquired.

2.4. Statistical analysis

Data were analyzed using Statistical Parametric Mapping (SPM12; www.fil.ion.ucl.ac.uk/spm/). With the exception of one participant whose head movement in the *z*-plane was approximately 3 mm in either direction, head movement was within 3 mm in all other cases. All functional volumes were spatially realigned to the first volume. A mean image created from realigned volumes was spatially normalized to the MNI EPI brain template using nonlinear basis functions. The derived spatial transformation was applied to the realigned T2* volumes and spatially smoothed with an 8-mm FWHM isotropic Gaussian kernel. Time series across each voxel were high-pass filtered with a cut-off of 128 sec, using cosine functions to remove section-specific low-frequency drifts in the BOLD signal. Condition effects at each voxel were estimated according to the general linear model, and regionally specific effects were compared using linear contrasts. The BOLD signal was modeled as a box-car, convolved with a canonical hemodynamic response function. We applied a cluster-level correction within SPM12 for determining statistical significance. Specifically, reported activations survived a voxel-level threshold of $p < .001$ (uncorrected for multiple comparisons) and a cluster-level threshold of $p < .05$, corrected for multiple comparisons (whole-brain FWE: Family-wise error). This statistical threshold was made identical to Vartanian et al. (2018) to enable us to compare our findings based on our two-response paradigm to earlier results based on the standard single-response paradigm.

Using an event-related design, in the first level, we specified regressors corresponding to the following time points in the problem structure (see Figure 1): (1) fixation point, (2) the groups in question, (3) stereotype information, (4) base rates, (5) prompt, (6) motor response for initial response, and (7) motor response for final response. In addition, the RTs associated with each motor response (i.e., initial and final) were included in the model as parameters and modeled out of the analyses by assigning a value of 0 to their respective regressor in subsequent analyses. All reported neural analyses are based on the prompt time points for the initial and final responses (i.e., last two slides in Figure 1). Importantly, prompt problems for initial and final responses were in turn separated into four separate regressors based on performance as follows: conflict (correct), conflict (incorrect), non-conflict (correct), non-conflict (incorrect).

3. Results

3.1. Behavioural results

A Conflict (2: conflict, non-conflict) × Response (2: initial, final) repeated-measures ANOVA on the accuracy data demonstrated a main effect for Conflict, $F(1, 21) = 25.56$, $p < .001$, partial eta-squared = .55, but there was no main effect for Response ($F[1, 21] = .67$, $p = .42$), or a Conflict × Response interaction ($F[1, 21] = .05$, $p = .83$) (Figure 2). Specifically, accuracy was lower for conflict than non-conflict problems both for initial ($t[21] = -5.08$, $p < .001$, $d = -1.15$) and final ($t[21] = -4.97$, $p < .001$, $d = -1.11$) responses.[1] These accuracy rates are consistent with earlier base-rate neglect findings involving the standard single-response administration of the same paradigm (see Vartanian et al., 2018), now extended to initial (conflict = 54.36%, $SD = 38.59$ vs. non-conflict =

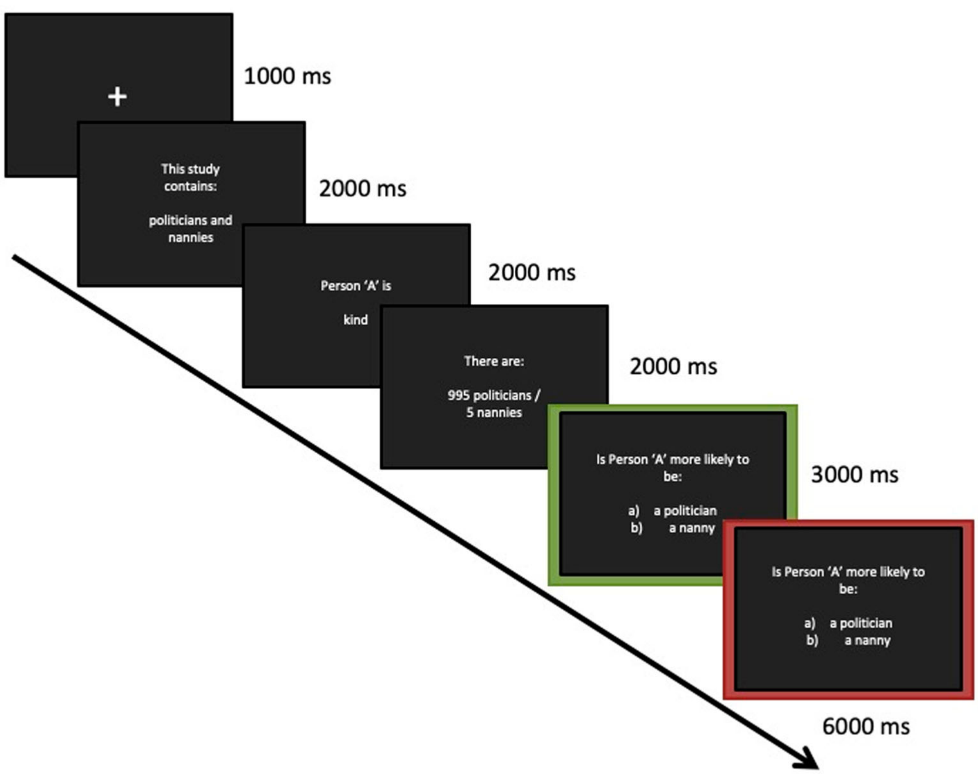

Figure 1. Procedure for the base rate task. [To view this figure in colour, please see the online version of this journal.]
Notes: Adapted with permission from Pennycook et al. (2014).

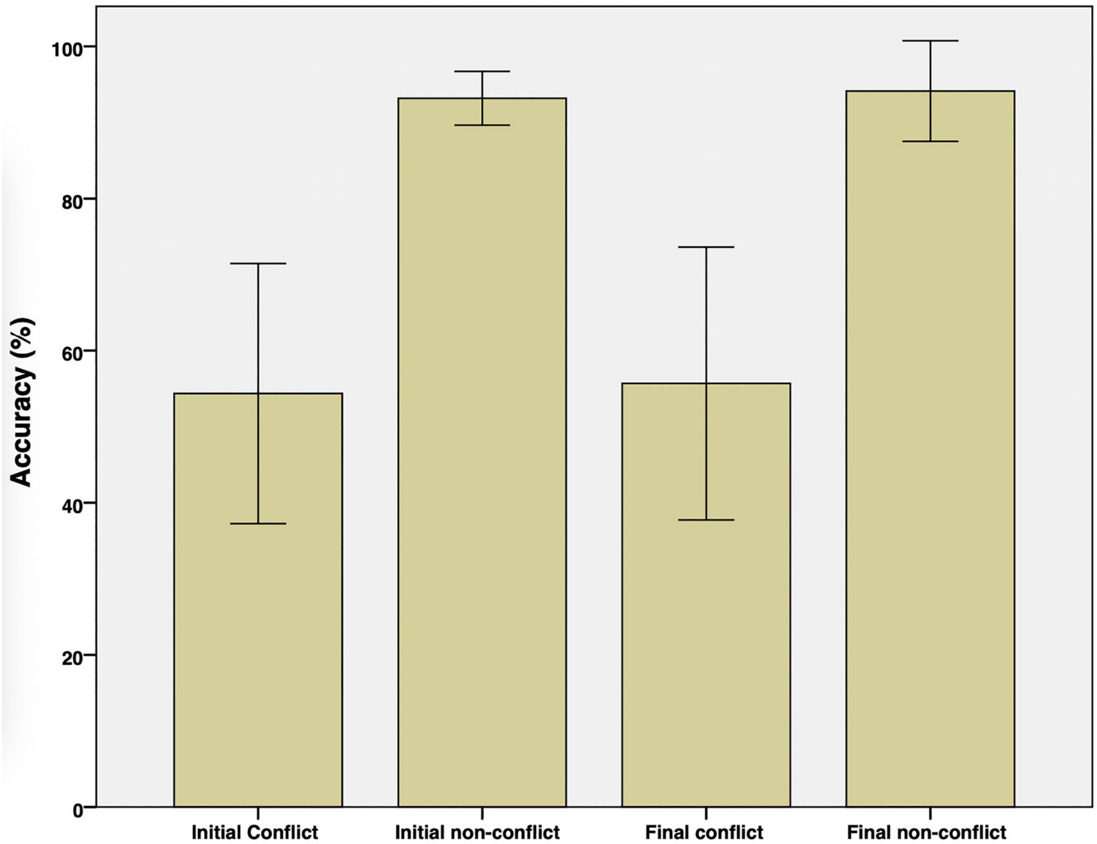

Figure 2. The effects of Response (initial, final) and Conflict (conflict, no conflict) on accuracy. [To view this figure in colour, please see the online version of this journal.]
Notes: Bars indicate 95% confidence intervals.

93.18%, $SD = 7.98$) and final (conflict = 55.68%, $SD =$ 40.47 vs. non-conflict = 94.13%, $SD = 14.92$) responses. Despite the observation that the patterns of performance on both conflict and non-conflict problems were similar for initial and final responses, at a global level participants nevertheless reported a higher degree of confidence in their performance for final ($M = 90.96\%$, $SD = 13.80$) than initial ($M = 85.71\%$, $SD = 16.$) responses, $t(23) = 3.00$, $p = .006$, $d = .33$. However, it is important to note that we collected a global confidence measure following the completion of the task, rather than on a trial-by-trial basis.

Next, we conducted a Conflict (2: conflict, non-conflict) × Response (2: initial, final) repeated-measures ANOVA on the reaction time (RT) data involving accurate responses only. The results demonstrated a main effect for Response, $F(1, 21) = 10.77$, $p = .004$, partial eta-squared = .34, but there was no main effect for Conflict ($F[1, 21] = 3.34$, $p = .14$), or a Conflict × Response interaction ($F[1, 21] = .08$, $p = .78$) (Figure 3). Specifically, RT was longer for initial than final responses both for conflict ($t[21] = 2.33$, $p = .03$, $d = .58$) and non-conflict ($t[21] = 3.56$, $p = .001$, $d = .90$) problems.

To put these RT data into perspective, note that previous work with the exact same item set involving the standard single-response administration—in which reflection was not restricted (Vartanian et al., 2018)—indicated that the critical correct conflict responses took on average 1941ms ($SD = 620$). Average initial response RT for correct conflict responses in the current study was 1358 ms ($SD =$ 401), or on average about one SD faster than unrestricted responding. This establishes that participants were under considerable time pressure when giving their initial responses and respected the instruction to respond as fast as possible.

In line with previous work, we also contrasted RT for correct non-conflict responses and incorrect conflict responses. This index is often used as a behavioural measure of biased reasoners' conflict sensitivity (De Neys & Pennycook, 2019). For initial responses, RT was longer for incorrect conflict than correct non-conflict trials, $t(18) = 2.63$, $p = .017$, $d = .57$. In

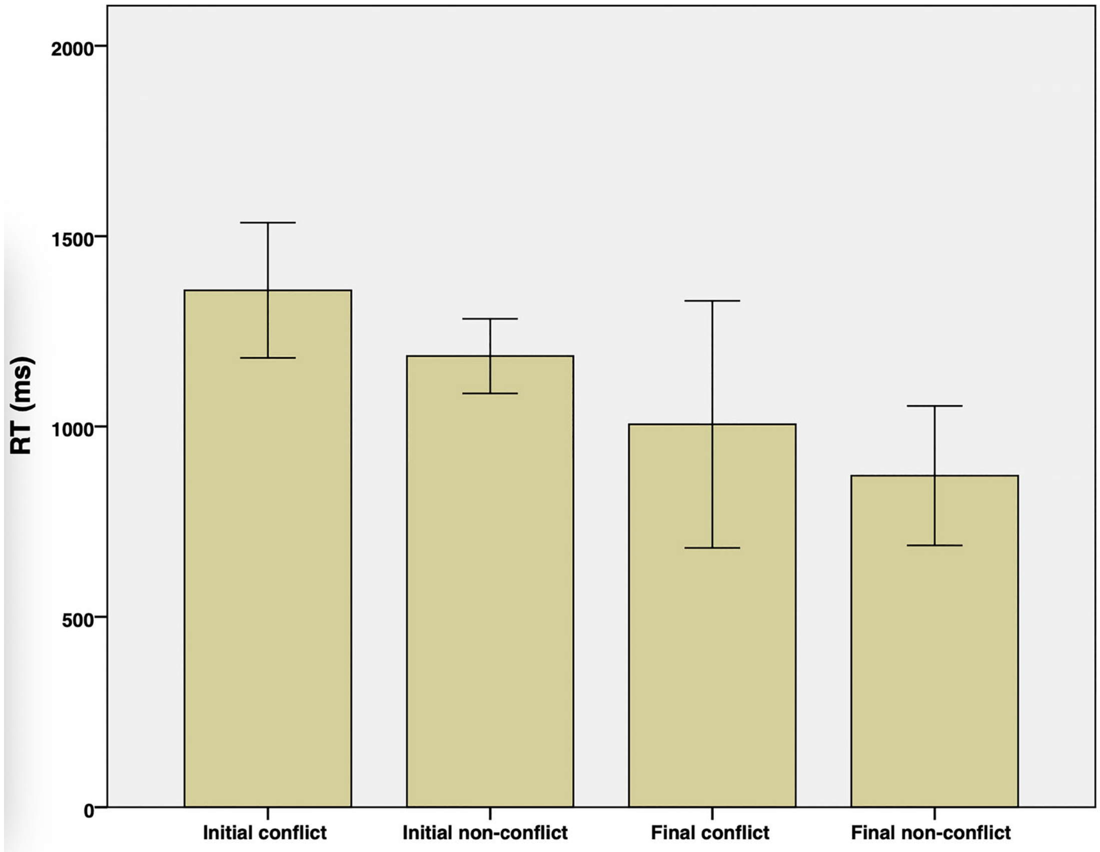

Figure 3. The effects of Response (initial, final) and Conflict (conflict, no conflict) on reaction time. [To view this figure in colour, please see the online version of this journal.]
Notes: Bars indicate 95% confidence intervals.

contrast, for final responses, this difference was in the same direction but not statistically significant, $t(17)$ = 2.63, $p = .18$, $d = .77$.

To explore in more detail whether the opportunity to make a final response had any impact on judgment, we conducted a change-of-direction analysis (Bago & De Neys, 2017) in which the frequency of responses given by all participants across all trials was binned into four categories: Correct-to-correct, correct-to-incorrect, incorrect-to-incorrect, and incorrect-to-correct (Figure 4). As can be seen, in only 4% of cases did longer deliberation change an initially biased response to a normatively correct response. Thus, by and large, participants did not change their initial responses when given the opportunity to do so.

As one reviewer suggested, here it is interesting to ask whether participants were more likely to change their final response if they had reasoned longer for the initial response. To test this idea, we conducted a binary logistic regression in which the final response (changed vs. unchanged) was regressed onto the RT for the initial response. The results demonstrated that longer RT for the initial response was associated with a greater likelihood of changing the final response, $ß = .001$, S.E. = .000, Wald = 65.91, $p < .001$. Next, we reran this analysis again, but now focusing only on the subsample of trials that were associated with non-normative responding in the initial stage. This more specific analysis is interesting for seeing whether one is more likely to revise one's choice if more effort had been put into the initial response. Here, too, the results demonstrated that longer RT for the initial response was associated with a greater likelihood of changing the final response, $ß = .001$, S.E. = .000, Wald = 18.39, $p < .001$.

Finally, we were interested in exploring individual differences, and their relationship with intuitive logic. Descriptive statistics associated with the individual-differences measures appear in Table 1. In turn, correlations between individual-differences and performance measures appear in Table 2. We began by first examining the distribution of

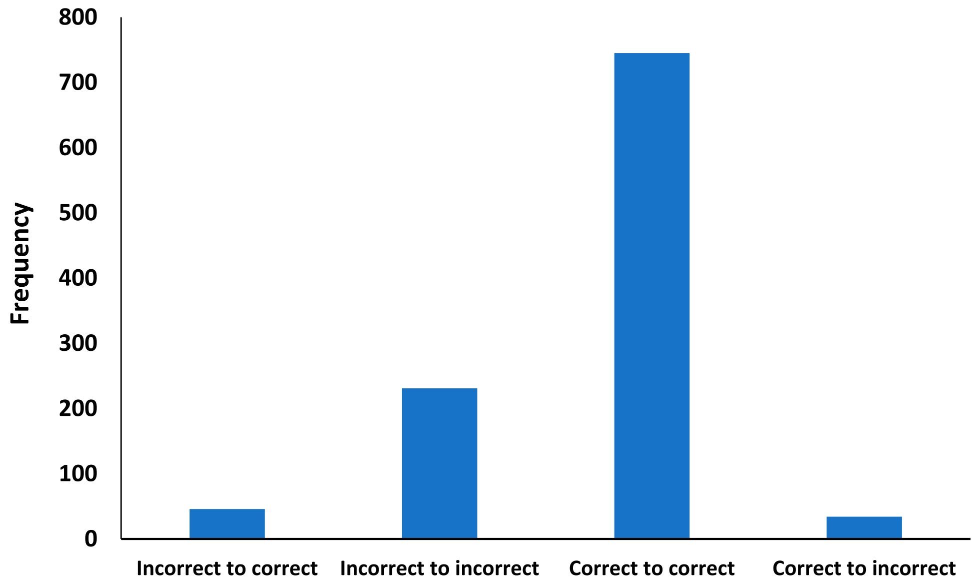

Figure 4. Direction-of-change data from initial to final responses. [To view this figure in colour, please see the online version of this journal.]

performance (i.e., accuracy) on conflict problems, separately for initial and final responses. Indeed, the distributions for both initial (Kolmogorov-Smirnov = .20, $p = .018$) and final (Kolmogorov-Smirnov = .22, $p = .007$) responses deviated from normal, and were bimodal in both cases. In other words, for both initial and final responses, there appear to be two groups of participants: Some who excel in providing the normative response, and some who do not. To examine whether it is the same participants who excelled under both conditions on conflict problems, we computed the zero-order Pearson correlation coefficient between accuracy for initial and final responses (averaged across each participant), which was near unity, $r(22)$ .99, $p < 001$. This suggests that it was the same participants who excelled in performance on conflict problems, regardless of time pressure.

Next, to better understand which cognitive ability and thinking style measures were correlated with our outcome measures of interest, we examined their correlations with performance on conflict and non-conflict problems for initial and final responses. As can be seen, our measure of crystallized intelligence (i.e., Vocabulary) was correlated positively with performance on conflict and non-conflict problems for both initial and final responses (Table 2). In contrast, our measure of fluid intelligence (i.e., Block Patterns) and CRT scores were correlated positively with performance on conflict but not non-conflict problems for both initial and final responses (Table 2). This suggest that the ability to think logically in novel situations (i.e., fluid intelligence) as well as the

Table 1. Descriptive statistics for individual-differences measures.

Measure	Mean
CRT (% accuracy)	52.98 (32.76)
Shipley-2 Vocabulary (standardized score)	107.25 (10.78)
Shipley-2 Block Patterns (standardized score)	105.88 (13.81)
Simple verbal WM span (STM)	9.57 (1.31)
Simple matrix WM span (STM)	8.90 (1.96)
Stroop (inhibition) in milliseconds	62.90 (60.90)
BFAS Conscientiousness	3.73 (.40)

Notes: Standard deviations appear in parentheses. CRT = Cognitive Reflection Test; BFAS = Big Five Aspect Scales; WM = working memory, STM = short-term memory. For Stroop the value reflects mean RT between the correct identification of incongruent trials vs. the correct identification of congruent trials (see text). The span measures were calculated based on partial-credit unit scoring (see Conway et al., 2005).

Table 2. Intercorrelations between individual-differences and performance measures.

	Initial conflict	Initial non-conflict	Final conflict	Final non-conflict
CRT	.45*	.39	.44*	.32
Shipley-2 Vocabulary	.54**	.51*	.52*	.49*
Shipley-2 Block Patterns	.50*	.35	.51*	.29*
Verbal span	.23	.39	.20	.27
Matrix span	.27	.23	.31	.31
Stroop	.19	.23	.20	.16
BFAS Conscientiousness	−.17	−.24	−.24	−.18

Notes: Values in the table reflect zero-order Pearson correlations. CRT = Cognitive Reflection Test; BFAS = Big Five Aspect Scales. * = $p < .05$, ** = $p < .01$.

ability to think reflectively (i.e., CRT) are related uniquely to judgment in relation to conflict problems regardless of time pressure. Here it is important to note that CRT is assumed to assess one's tendency to override a prepotent response that is incorrect, and to engage in further reflection that leads to the correct response. In this sense, it is perceived as the quintessential measure for one's willingness and ability to exert effort to optimize judgment and decision making. CRT has been shown to be a reliable predictor of rational performance in a wide host of tasks (e.g., expected-value choices/gambles, temporal discounting, framing, conjunction fallacy, maximizing strategies on probabilistic prediction tasks, endorsement of profit maximizing strategies, avoidance of the illusion of explanatory depth, non-superstitious thinking, performance calibration, and general numeracy) (see Toplak et al., 2014). It has also shown similar predictive power based on a composite measure that includes fifteen separate rational thinking tasks from many different domains in the heuristics and biases literature (Toplak et al., 2011). In turn, Block Patterns can be considered a measure of visuospatial fluid intelligence, and intelligence tests have also been shown to predict rational thinking performance in many contexts (Stanovich & West, 1998, 2000, 2008). As such, the present findings add to a large body of evidence that supports the utility of CRT and intelligence measures as predictors of rational thinking in many contexts. Moreover, in line with recent findings, the correlation with the initial (intuitive) conflict performance indicates that in addition to the tendency to reflectively correct an incorrect intuition it also may track the ability to generate correct intuitions in the absence of reflection (e.g., Bago & De Neys, 2019b; Raoelison et al., 2020)

3.2. Neural results

We conducted a series of analyses to explore the neural correlates of logical intuitions. Unless stated otherwise, in each case the reverse contrast was also carried out but did not lead to any significant area of activation. The specific number of trials that fell within each cell of the design across all 24 participants—broken down by accuracy—are presented in Table 3. Although our neural hypotheses were not based on interactions, for the sake of consistency

Table 3. The specific number of trials that fell within each cell of the design across all 24 participants—broken down by accuracy.

Condition	Frequency
Initial Response	1056
Conflict	528
Correct	287
Incorrect	241
Non-conflict	528
Correct	492
Incorrect	36
Final Response	1056
Conflict	528
Correct	294
Incorrect	234
Non-conflict	528
Correct	497
Incorrect	31

with the flow of our behavioural analyses, we conducted the Conflict (2: conflict, non-conflict) × Response (2: initial, final) repeated-measures ANOVA in SPM12. This analysis did not yield any significant activation.

We began by comparing activation for conflict vs. non-conflict problems, separately for initial and final responses. Previous work had indicated that frontal control regions are specifically brought online for conflict reasoning problems (e.g., De Neys et al., 2008; Goel & Dolan, 2003; Goel et al., 2017; Noveck et al., 2004; Stollstorff et al., 2012; Tsujii et al., 2010; Vartanian et al., 2018). Our current analysis revealed that for initial and final responses, there was no difference in brain activation between conflict vs. non-conflict problems. Next, we proceeded to examine the same contrast, but separately for correct and incorrect responses. Examining correct and incorrect responses separately was done not only because this is a standard analytical approach in behavioural conflict detection studies on reasoning (e.g., De Neys & Pennycook, 2019), but more importantly because *a priori*, from a theoretical perspective, we were particularly interested to examine whether there would be evidence of conflict sensitivity associated with incorrect responding. This is because theoretically, it is generally not disputed that correct responders will show conflict sensitivity. Indeed, many models assume that detecting conflict is a necessary precondition in overcoming it in favour of normative responding. In contrast, it is of interest to see whether incorrect responders, despite poor performance, nevertheless exhibit conflict sensitivity. Toward this end, we broke down this contrast by

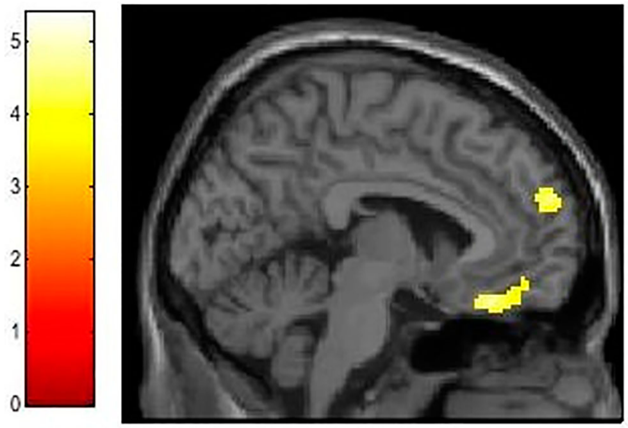

Figure 5. Initial incorrect conflict trials vs. non-conflict trials. [To view this figure in colour, please see the online version of this journal.]

Notes: Incorrect initial conflict trials activated the left medial frontal gyrus (BA 9) and the anterior cingulate cortex (BA 32) (see text and Table 4). SPM rendered into standard stereotactic space and superimposed onto sagittal MRI in standard space. Bar represents the corresponding *T* score.

the accuracy of conflict trials. Specifically, the comparison of initial incorrect conflict trials vs. non-conflict trials revealed relatively greater activation in left medial frontal gyrus (BA 9) and the ACC (BA 32) (Figure 5, Table 4). To examine whether this finding might have been affected by individual differences in performance on conflict trials, we reran this analysis but with average accuracy on conflict trials entered into the GLM as a covariate. As before, we observed significant activations in the same locations within the left medial frontal gyrus (BA 9) and the ACC (BA 32) (see Table 4). In contrast, the comparison of initial correct conflict trials vs. non-conflict trials did not reveal any significant area of activation. Next, as in Vartanian et al. (2018), we also ran these analyses with only correct non-conflict responses as the contrast group. The comparison of initial incorrect trials vs. correct non-conflict trials did not reveal any significant area of activation. Similarly, the comparison of initial correct conflict trials vs. correct non-conflict trials did not reveal any significant area of activation. For final responses, the comparisons of correct or incorrect conflict trials vs. (correct) non-conflict trials did not reveal any area of activation.

Next, we compared initial incorrect conflict trials to final incorrect conflict trials. This contrast revealed relatively greater neural activation in a distributed set of

Table 4. Regions activated by various contrasts of fMRI data.

Contrast	Region	BA	Coordinates	Cluster size	*p*
Initial conflict trials (incorrect) – initial non-conflict trials					
	Medial frontal gyrus	9	−6, 60, 24	173	.045
	Anterior cingulate cortex	32	−4, 34, −20	206	.023
Initial conflict trials (incorrect) – initial non-conflict trials (with conflict accuracy as covariate)					
	Medial frontal gyrus	9	−8, 58, 24	155	.050
	Anterior cingulate cortex	32	−4, 34, −20	169	.039
Initial conflict trials (incorrect) – final conflict trials (incorrect)					
	Inferior frontal gyrus	9	−40, 10, 32	444	<.001
	Precentral gyrus	6	44, 6, 30	142	.033
	Middle frontal gyrus	9	40, 34, 26	232	.003
	Anterior prefrontal cortex	10	34, 54, −4	128	.050
	Superior parietal lobule	7	28, −58, 60	258	.002
	Superior parietal lobule	7	−26, −62, 58	210	.005
	Middle occipital gyrus	19	40, −84, 6	510	<.001
	Inferior occipital gyrus	18	−36, −90, −8	207	.006
	Lingual gyrus	18	6, −72, 10	166	.017
Initial conflict trials (incorrect) – final conflict trials (incorrect) (with conflict accuracy as covariate)					
	Inferior frontal gyrus	9	−40, 10, 32	511	<.001
	Middle frontal gyrus	9	40, 34, 26	390	<.001
	Anterior prefrontal cortex	10	34, 54, −4	125	.050
	Superior parietal lobule	7	28, −58, 60	220	.004
	Superior parietal lobule	7	−26, −62, 58	184	.009
	Middle occipital gyrus	19	40, −84, 6	603	<.001
	Inferior occipital gyrus	18	−36, −90, −8	233	.003
	Lingual gyrus	18	12, −74, 8	146	.027
	Cerebellum	–	18, −84, −24	135	.037
Initial conflict trials (correct) – final conflict trials (correct)					
	Lingual gyrus	18	−6, −76, 8	375	<.001
Final conflict trials (incorrect) – final non-conflict trials (correct)					
	Medial/superior frontal gyrus	8	−10, 46, 46	362	.003

Notes. BA = Brodmann Area. All reported activations survived cluster-level intensity threshold of $p < .05$, corrected for multiple comparisons using the whole brain family-wise error (FWE) correction (see specific *p*-value above). Regions are designated using the MNI coordinates.

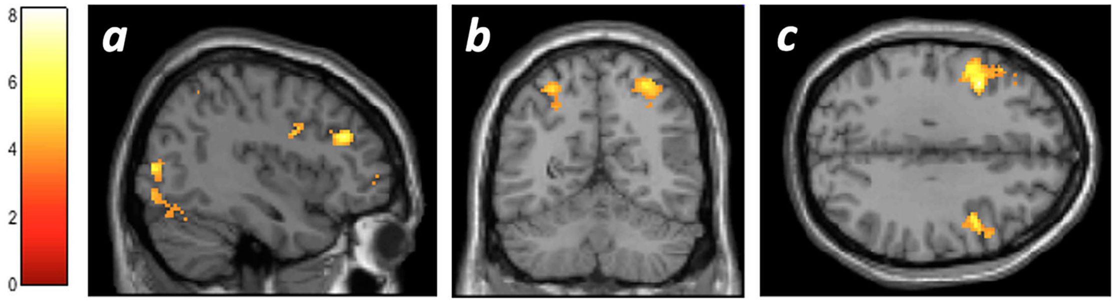

Figure 6. Initial (incorrect) conflict trials – final (incorrect) conflict trials. [To view this figure in colour, please see the online version of this journal.]

Notes: Initial (incorrect) conflict trials activated (a) middle occipital gyrus, middle frontal gyrus, precentral gyrus, and anterior prefrontal cortex, (b) bilateral superior parietal lobule, and (c) inferior frontal gyrus and precentral gyrus (see text and Table 4). SPM rendered into standard stereotactic space and superimposed onto (a) sagittal, (b) coronal, and (c) axial MRI in standard space. Bar represents the corresponding *T* score.

regions in the frontal, parietal, and occipital lobes (Figure 6, Table 4). To examine whether this finding might have been affected by individual differences in performance on conflict trials, we reran this analysis but with average accuracy on conflict trials entered into the GLM as a covariate. With the exception of the precentral gyrus, we observed activations in the same regions as before; in addition, we observed activation in the cerebellum (see Table 4). Next, we contrasted initial correct conflict trials to final correct conflict trials. This contrast revealed relatively greater activation exclusively in the lingual gyrus (Figure 7, Table 4). Then, we contrasted final incorrect conflict trials to final correct non-conflict trials, which revealed relatively greater activation exclusively in the medial/superior frontal gyrus (BA 8) (Figure 8, Table 4). Finally, we contrasted final correct conflict trials to final correct non-conflict trials, which did not reveal any statistically significant difference.

4. Discussion

We conducted our experiment to test three hypotheses. First, we hypothesized that participants would exhibit logical intuitions reflected by similar rates of correct, normative responses regardless of time pressure, and a low rate of response switches following deliberation. This hypothesis was confirmed. Indeed, the pattern of normative responding was invariant to time pressure (Figure 2), and when given the opportunity to reconsider their choices made in the initial response, by and large participants did not

Figure 7. Initial (correct) conflict trials – final (correct) conflict trials. [To view this figure in colour, please see the online version of this journal.]

Notes: Initial correct conflict trials activated the left lingual gyrus (BA 18) (see text and Table 4). SPM rendered into standard stereotactic space and superimposed onto sagittal MRI in standard space. Bar represents the corresponding *T* score.

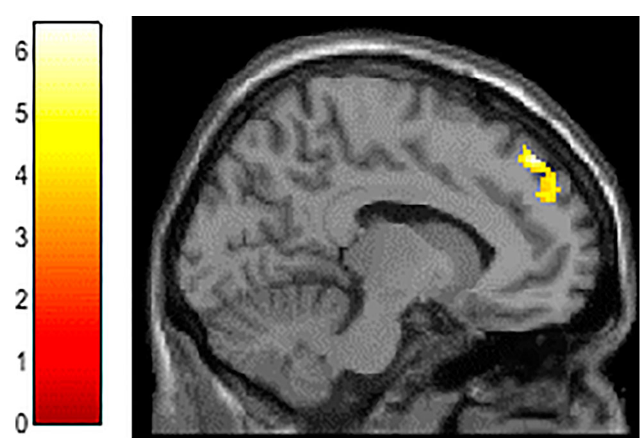

Figure 8. Final (incorrect) conflict trials – final (correct) non-conflict trials. [To view this figure in colour, please see the online version of this journal.]

Notes: Final (incorrect) conflict trials activated the left medial/superior frontal gyrus (BA 8) (see text and Table 4). SPM rendered into standard stereotactic space and superimposed onto sagittal MRI in standard space. Bar represents the corresponding *T* score.

change their minds in the final response stage (Figure 4). Critically, in only 4% of cases did longer deliberation change an initially biased response to a normatively correct response. Also, final RTs were not longer than initial RTs, indicating further that people do not tend to deliberate much in the task when given the opportunity to do so. This reluctance to alter an initially incorrect choice despite further opportunity may be an example of inaction or decision inertia (e.g., Anderson, 2003). As noted by Anderson (2003), this tendency can be due to a number of different factors including cost– benefit calculations, anticipated regret, and selection difficulty, although the specific mechanism that led to not switching in our case is not entirely clear. For example, RT was longer for initial than final responses both for conflict and non-conflict problems (Figure 3). At the same time, participants were more likely to switch their final responses if they had deliberated longer toward their initial intuitively-cued but biased response. It is possible that a longer deliberation time might serve as an internal cue that one's initial response might be non-normative, thus prompting revision. This possibility can be studied in the future. Nevertheless, participants also reported greater confidence in their final than initial choices, suggesting that at a metacognitive level the opportunity for further reflection can increase one's assessment of one's own level of performance (Thompson et al., 2011).

Second, we hypothesized that superior performance on conflict problems, under time pressure or otherwise, would be correlated with individual differences in cognitive abilities and thinking styles. This hypothesis was also confirmed. Specifically, we found that fluid intelligence and CRT scores were correlated with performance on conflict but not on non-conflict problems for both initial and final responses. This result is consistent with a large body of literature demonstrating that cognitive ability and the tendency to think reflectively are related to performance in a range of heuristics and biases tasks, including base-rate neglect (e.g., Toplak et al., 2011, 2014; Vartanian et al., 2018). However, this result is particularly interesting with respect to CRT in relation to *initial* responses because it suggests that the tendency to think reflectively is related to and can contribute to fast logic—consistent with the idea that despite their speed, logical intuitions can nevertheless be based on a thinking style that supports logical thought more generally. As noted by Thompson et al. (2018; see also Raoelison et al., 2020), better thinkers might have better intuitions which can be deployed quickly.

Although our study was not designed to explore the neural correlates of individual differences in relation to brain activation, for exploratory purposes we conducted multiple regression analyses in SPM12 to probe the correlation between individual-differences measures and brain activation on initial and final conflict and non-conflict problems. Our results demonstrated that Shipley Verbal scores were correlated with activation in the cingulate gyrus, bordering on the caudate, for both correct conflict (BA = 32, $x = -12$, $y = 26$, $z = 12$, $T = 5.59$, $k_E = 230$) and non-conflict (BA = 32, $x = -16$, $y = 28$, $z = 24$, $T = 4.47$, $k_E = 163$) final responses. This region has been associated with semantic retrieval, as well as the interaction of working memory with the semantic system (Binder et al., 2009; Deldar et al., 2020). No other correlation was statistically significant. This finding suggests that higher verbal intelligence and the ability to recruit relevant semantic information could play a role in correct responding, but only under conditions in which time pressure is not an issue. From a mechanistic perspective, the association of verbal intelligence with final but not initial responding might be driven by the time course required to access and incorporate semantic information into the judgment and decision-making stream—a process that could require a longer time window for full deployment.

Third, from a neural perspective, we had hypothesized that participants would already exhibit conflict sensitivity when giving quick, initial responses, reflected by brain activation in regions of the brain that underlie error detection, attention and cognitive control on initial responses (compared to final responses). If so, this would indicate that neural regions involved in detecting conflict can be engaged to support cognition under time pressure. This hypothesis was also confirmed (Table 4). Specifically, the contrast of initial incorrect conflict trials to final incorrect conflict trials revealed activation in left inferior frontal gyrus (BA 9), right middle frontal gyrus (BA 9), right anterior frontal gyrus (BA 10), right precentral gyrus (BA 6), bilateral superior parietal lobule (BA 7), right middle occipital gyrus (BA 19) and left inferior occipital gyrus (BA 18). This set of regions in the frontal and parietal lobes has been implicated heavily in cognitive control and executive functions (Cole & Schneider, 2007; Dosenbach et al., 2006; Miller, 2000; Miller & Cohen, 2001; Ridderinkhof et al., 2004b; see also Lenartowicz et al., 2010), as well as various aspects of visual attention (Kanwisher & Wojciulik, 2000; Kastner & Ungerleider, 2000; Wager et al., 2004; Vartanian et al., 2018). Previous imaging work in the reasoning field in which participants' reflection was not restricted has typically also pointed to activations in these regions (e.g., De Martino et al., 2006; De Neys et al., 2008; Goel et al., 2000; Goel & Dolan, 2003; Houdé et al., 2011; Prado et al., 2011; Prado & Noveck, 2007; Stollstorff et al., 2012; Tsujii & Watanabe, 2010; Tsujii et al., 2010; Vartanian et al., 2018). This suggests that regions that support processes that one would consider necessary for sensitivity to conflict—including attention, inhibition, and interference control—can come online rapidly (i.e., "intuitively"), and that this is the case despite having made a biased response.

On the other hand, the contrast of initial correct conflict trials to final correct conflict trials did not reveal activation in any of these cognitive control or executive functions regions, instead activating only the lingual gyrus. How can it be that generating normative responses accurately and quickly does not seem to engage a large set of frontal and parietal brain regions in comparison to making the same responses deliberately? This seemingly counterintuitive observation is consistent with the idea that throughout development, most adult thinkers come to largely automatize the principles associated with logical and normative thinking, perhaps through exposure to education. Specifically, De Neys and Pennycook (2019) have argued that people's intuitive logical knowledge emerges from a learning process in which over time, key principles have been practiced to a point where they become automatic (De Neys, 2012). The general idea that initially effortful processes can become automatized over time is not new (Kahneman, 2011; Shiffrin & Schneider, 1977; Stanovich, 2018), although its application to reasoning is relatively novel. As such, one would indeed expect that performance in accordance with norms under time pressure may activate largely the same set of regions that would be active given more time, since in both cases the process might have been largely automatized due to practice and learning. This is also consistent with experimental data demonstrating that good thinkers have good intuitions, and as such have the ability to think logically, quickly, and without deliberation (Raoelison et al., 2020; Thompson et al., 2018). Note that this conclusion also fits with recent neuroimaging work in which adolescents' and adults' performance on the ratio bias task was contrasted (Mevel et al., 2019). Although Mevel et al. found that adults behaviourally outperformed adolescents, such correct responding was associated with *less* frontal activation in cognitive control regions for adults than adolescents. In line with the present findings, Mevel et al. suggested that because of better intuiting, correct responding might become less demanding for soundly reasoning adults.

In addition to the contrast between initial and final conflict responses, we also compared activation for conflict and non-conflict problems. Previous work has indicated that frontal control regions are specifically brought online for conflict reasoning problems (e.g., De Neys et al., 2008; Goel & Dolan, 2003; Stollstorff et al., 2012; Tsujii et al., 2010; Vartanian et al., 2018). Here we did not observe greater activation in conflict vs. non-conflict trials in the ACC. The reason for this is unclear, although one possibility might be low statistical power. For example, we did observe activation in the left medial frontal gyrus that did not reach statistical significance ($p = .08$), and there is evidence to suggest that the medial frontal cortex is involved in error monitoring and response selection, as well as the context within which choices are made (see Rushworth et al., 2007). However, a closer exploration indicated that—at least for incorrect initial conflict responses—there is in fact evidence for conflict sensitivity. For example, the comparison of initial incorrect conflict trials vs. non-conflict trials revealed significant activation in left medial frontal gyrus (BA 9) and ACC (BA 32) (Figure 5, Table 4). The sensitivity of the medial frontal gyrus and the ACC to prediction errors has been well established through meta-analysis (Garrison et al., 2013), and their activations in relation to incorrect conflict trials suggests a similar sensitivity on part of the participants. Critically, however, none of these more subtle conflict activations were observed for final incorrect conflict responses, or correct responders' initial or final conflict responses. This pattern supports our main finding in the initial vs. final response contrasts: Incorrect responders recruit frontal control regions when faced with initial but not final conflict problems.

This indicates that the conflict-related activations we observed for incorrect responders mainly result from quick, intuitive processing rather than slow deliberation. Furthermore, the lack of specific conflict-related neural activation for correct initial and final responses again suggests that correct responders simply face little conflict and do not need to recruit cognitive control—either fast or slow.

Returning to the theme of individual differences, it is possible that early on, some participants might have selected and applied a rule (e.g., base rates) consistently throughout the remainder of the task. In this sense, the consistent and strategic application of a rule could itself be viewed as the representation of an intuitive/automated process. It is important to note that this possibility is not inconsistent with our account. For example, our neural activations demonstrate that *correct* responders are not exhibiting high levels of conflict/control-related activation. As noted elsewhere, it is possible that sound reasoners are more likely to focus exclusively on the base rates, thereby experiencing less conflict. To demonstrate this visually based on the behavioural data, we have plotted the histograms of response accuracy on the base rate task for the four trial types (Figure 9). The bimodal shapes of the distributions associated with conflict trials suggests that sound reasoners might in fact be attending to base rates only. Nevertheless, lower accuracy on conflict vs. non-conflict trials suggests that some participants are nevertheless affected by conflict between base rates and individuating information, and hence, it is unlikely that they simply discard the latter and engage exclusively in blind rule application. The key point is that our neural data point to conflict sensitivity in the case of *incorrect* conflict responders, and that paired with the classic behavioural RT finding that indicates longer RT for incorrect conflict over non-conflict trials, this demonstrates that incorrect responders are experiencing conflict due to processing both pieces of information (i.e., base rates and individuating information).

To our knowledge, the present study was the first to implement a two-response paradigm in a

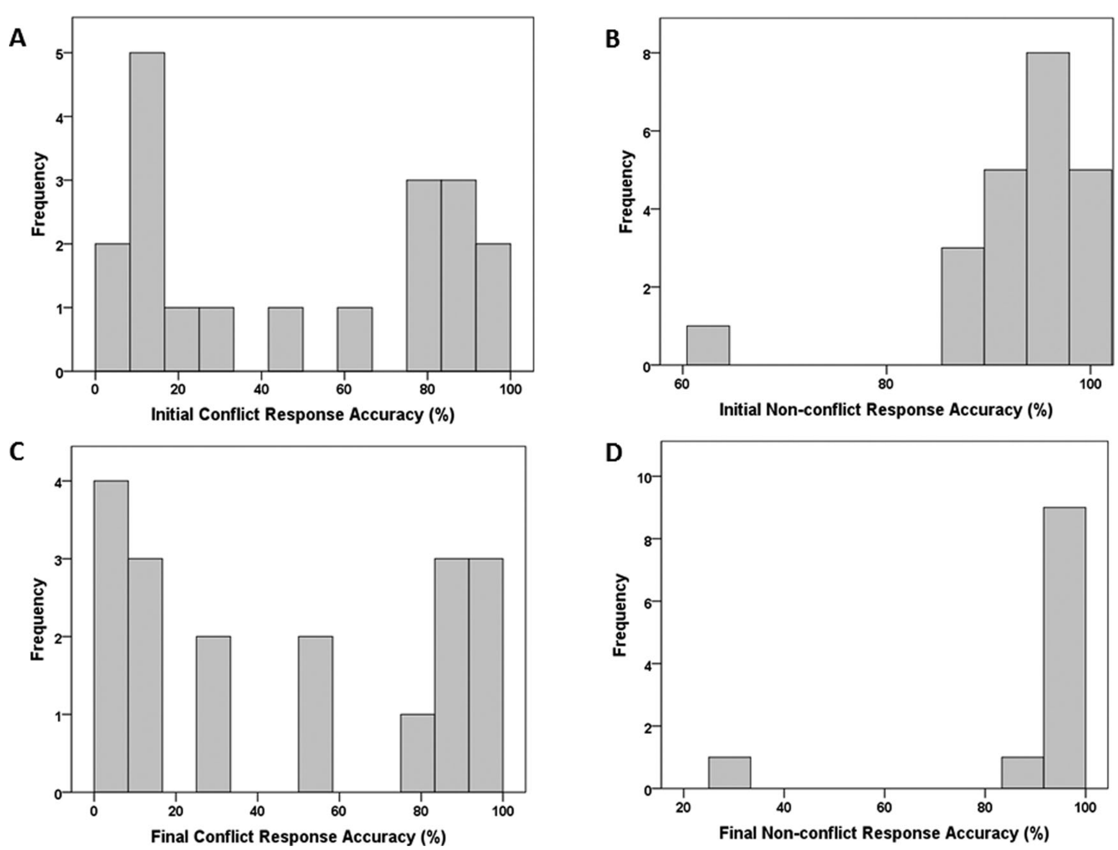

Figure 9. Histograms of response accuracy on the base rate task for the four trial types (i.e., initial conflict, initial non-conflict, final conflict, and final non-conflict trials).

neuroimaging study involving fMRI. Hence, results must obviously be interpreted with some degree of caution. It will be important to replicate and extend the work in future studies. For example, the present study focused on the base-rate neglect task since the problem has been a paradigmatic test-bed in the (behavioural) literature (e.g., De Neys & Glumicic, 2008; Pennycook et al., 2015). However, by now, behavioural two-response studies have been validated with a wide range of other tasks as well (De Neys, 2017; De Neys & Pennycook, 2019). Such cross-task generalization will be important for future neuroimaging work too.

To avoid confusion, we also would like to address some possible misconceptions. On one hand, one might argue that our instructions were not restrictive enough. Participants might not have been time-pressured, and already deliberated during the initial response phase. As we noted, this argument is countered by the observation that participants' initial responses were given considerably faster than under standard single-response administration in which they are allowed to deliberate on this identical problem set (see Vartanian et al., 2018). This indicates that participants respected the instruction to respond as fast as possible, and that their reasoning relied more on intuitive rather than deliberate processing. Nevertheless, given that the dual-process framework does not specify a clear threshold that allows us to classify a process as uniquely intuitive or deliberate, we can never be completely sure that participants relied on pure intuitive reasoning (e.g., Bago & De Neys, 2019b; De Neys, 2021). What is important here is that our findings indicate that critical control regions can come online quickly under conditions that minimize possible deliberation.

On the other hand, one might argue that our instructions were not stimulating enough. Indeed, participants were allowed to deliberate in our final response stage but—as in traditional single-response studies—this does not imply that they effectively did so. Hence, we do not contest that if participants were stimulated or forced to deliberate, this might have resulted in different results. More generally, the present findings should not be taken to imply that people cannot deliberate or that such deliberation would not be associated with specific neural activations. Our point is simply that this is not spontaneously observed, and that critical conflict-related activation in conflict detection and control regions is already observed under time pressure when participants are forced to make quick judgments.

In addition, our two-response paradigm is predicated on the assumption that intuitive responding occurs under time pressure. Consistent with this assumption, the analysis of our RT data demonstrated that the average RT for initial correct conflict responses in the current study was about one SD faster than unrestricted responding in our previous work with the exact same item set involving the standard single-response administration where reflection was not restricted (Vartanian et al., 2018). Note that in behavioural studies, comparison with the unrestricted response time condition is considered the gold standard for ascertaining the involvement of intuitive processes (e.g., Bago et al., 2021; Lawson et al., 2020). In other words, our results support the argument that participants responded quickly. In addition, it is not necessary to observe a speed-accuracy tradeoff when comparing restricted to unrestricted trials if participants already mainly rely on intuitive processing in the unrestricted response condition as well. Indeed, average performance under the unrestricted condition for conflict trials was around chance level (i.e., approximately 50%), a level of performance that would be difficult to degrade further under time pressure. As such, we believe that the experimental design did in fact propel participants to respond intuitively due to the restricted time allowance associated with the initial response window (see also Bago & De Neys, 2017; Newman et al., 2017).

4.1. Neuropsychological studies

In closing, it is important to discuss the implications of our work for neuropsychological studies involving heuristics and biases tasks that pit intuitions and/or beliefs against rational norms. In his review of the cognitive neuroscience of reasoning, Goel (2007) noted that one of the oldest findings in the literature is the observation that people perform better on reasoning tasks when the logical conclusion is consistent rather than inconsistent with their beliefs. As such, it is important to ask which neural systems enable one to detect and resolve inconsistencies between logic and beliefs in the service of rational thought. There are substantial lesion and patient data to demonstrate that persons with lesions to right lateral prefrontal cortex perform poorly on such tasks, suggesting that this region plays a critical role in this process (for review see Goel, 2021). In turn, neuroimaging data based on base-rate tasks have shown a dissociation between conflict detection vs. resolution, suggesting that a neural system encompassing the dorsomedial prefrontal cortex including ACC might be sensitive to the presence of conflict (i.e., conflict detection), whereas a neural system encompassing the right lateral prefrontal cortex might be necessary to resolve the conflict in order to respond logically (see De Neys et al., 2008; see also Vartanian et al., 2018). In terms of advancing this work, it would be of great value to test the specific contributions of dorsomedial and right lateral prefrontal cortex to the detection vs. resolution of conflict in patients with focal brain lesions. Such data would be useful for determining whether the dorsomedial prefrontal cortex including ACC plays a necessary role in logical intuitions due to its role in conflict detection.

4.2. Limitations

There are a few methodological and analytical limitations associated with our study that must be highlighted. First, in terms of statistical power, when trials were broken down by accuracy, the cells in our design did not contain the same number of trials (Table 3). However, it is a reliable statistical and empirical feature of the base-rate neglect task we adopted that on average participants' accuracy on conflict vs. non-conflict trials hovers around 50% and 100% respectively, and as such a comparison of correct vs. incorrect conflict trials will always include fewer trials in the former category for conflict vs. non-conflict trials *regardless of how many sessions of the task are administered*. Second, because there was no temporal gap between the initial and final responses in our design (Figure 1), this likely reduced our ability to isolate the BOLD signal uniquely associated with each type of responses. Unfortunately, however, it was not possible to add a temporal gap between initial and final responses because the entire experimental paradigm rests on a tight sequencing of initial and final responses, as has been done in all previous behavioural work involving the two-response paradigm. This is because the introduction of a temporal separation opens the door to the possibility that the participant might engage in reasoning during that period—thus undermining the entire paradigm. The reliability of our neural findings can be tested in the future by manipulating cognitive load rather than time pressure, which should theoretically have the same resource-limiting effect on the reasoning machinery (see Bago & De Neys, 2017; Newman et al., 2017). Similarly, lesion and patient data can serve a similar purpose by determining whether specific regions that have been identified in this and similar neuroimaging studies are necessary for instantiating logical intuitions.

5. Conclusion

Our results add to a growing body of literature suggesting that the application of norms and logic to thinking need not necessarily be an effortful process, but can occur intuitively. They also demonstrate that cognitive abilities and thinking styles that have been shown to be correlated with normative performance in the face of conflict under slow conditions are also correlated with normative performance in the face of conflict under fast conditions. Finally, at the neural level, a set of regions in the frontal and parietal lobes that is involved in cognitive control, executive functions and attention is activated relatively more even when participants make biased judgments intuitively, suggesting that this set of regions might be sensitive to intuitive conflict detection. Our results are inconsistent with theories that posit that biased responding necessarily occurs outside of the window of awareness. Rather, they suggest that even in cases when logical errors are made, there might be some early awareness of conflict.

Note

1. Effect sizes for ANOVAs (partial eta-squared) are computed automatically in SPSS. Effect sizes for t-tests were calculated using an online calculator (Lenhard & Lenhard, 2016).

Acknowledgments

We would like to thank Gordon Pennycook for sharing with us the original E-Prime program and experimental material for the base rate task, which we modified for this experiment to implement the 2-response paradigm. We would like to also thank Joy Williams for her assistance in data collection involving MRI, and Nicole Herz for her help in the literature review.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by funding from Canada's Department of National Defence (DND) (02ac: Human Systems Integration).

ORCID

Oshin Vartanian http://orcid.org/0000-0002-4777-2513

References

Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin*, *129*(1), 139–167. https://doi.org/10.1037/0033-2909.129.1.139

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, *6*(2), 115–116. https://doi.org/10.1038/nn1003

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177. https://doi.org/10.1016/j.tics.2004.02.010

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*, *18*(4), 177–185. https://doi.org/10.1016/j.tics.2013.12.003

Bago, B., Bonnefon, J. F., & De Neys, W. (2021). Intuition rather than deliberation determines selfish and prosocial choices. *Journal of Experimental Psychology: General*, *150*(6), 1081–1094. https://doi.org/10.1037/xge0000968

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109. https://doi.org/10.1016/j.cognition.2016.10.014

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

Bago, B., & De Neys, W. (2019b). The smart System 1: Evidence for the intuitive nature of correct responding on the bat-andball problem. *Thinking & Reasoning*, *25*(3), 257–299. https://doi.org/10.1080/13546783.2018.1507949

Bago, B., Frey, D., Vidal, J., Houdé, O., Borst, G., & De Neys, W. (2018). Fast and slow thinking: Electrophysiological evidence for early conflict sensitivity. *Neuropsychologia*, *117*, 483–490. https://doi.org/10.1016/j.neuropsychologia.2018.07.017

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

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Cole, M. W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, *37*(1), 343–360. https://doi.org/10.1016/j.neuroimage.2007.03.071

Conway, A. R., 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

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9

Deldar, Z., Gevers-Montoro, C., Khatibi, A., & Ghazi-Saidi, L. (2020). The interaction between language and working memory: a systematic review of fMRI studies in the past two decades. *AIMS Neuroscience*, *8*(1), 1–32. https://doi.org/10.3934/Neuroscience.2021001

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684–687. https://doi.org/10.1126/science.1128356

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. (2017). *Dual process theory 2.0*. Routledge.

De Neys, W. (2021). On dual and single process models of thinking. *Perspectives on Psychological Science*, *16*(6), 1412–1427. https://doi.org/10.1177/1745691620964172

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*(3), 1248–1299. https://doi.org/10.1016/j.cognition.2007.06.002

De Neys, W., & Pennycook, G. (2019). Logic, fast and slow: Advances in dual-process theorizing. *Current Directions in Psychological Science*, *28*(5), 503–509. https://doi.org/10.1177/0963721419855658

De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*(5), 483–489. https://doi.org/10.1111/j.1467-9280.2008.02113.x

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big five. *Journal of Personality and Social Psychology*, *93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., & Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron*, *50*(5), 799–812. https://doi.org/10.1016/j.neuron.2006.04.031

Evans, J. S. B., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, *11*(4), 382–389. https://doi.org/10.1080/13546780542000005

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Evans, J. St. B. (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

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

Frey, D., Johnson, E. D., & De Neys, W. (2018). Individual differences in conflict detection during reasoning. *Quarterly Journal of Experimental Psychology*, *71*(5), 1188–1208. https://doi.org/10.1080/17470218.2017.1313283

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *37*(7), 1297–1310. https://doi.org/10.1016/j.neubiorev.2013.03.023

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 513–525. https://doi.org/10.1037/0096-1523.14.3.513

Goel, V. (2007). Anatomy of deductive reasoning. *Trends in Cognitive Sciences*, *11*(10), 435–441. https://doi.org/10.1016/j.tics.2007.09.003

Goel, V. (2021). Rationality and the brain. In M. Knauff & W. Spohn (Eds.), *Handbook of rationality*. MIT Press.

Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *NeuroImage*, *12*(5), 504–514. https://doi.org/10.1006/nimg.2000.0636

Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition*, *87*(1), B11–B22. https://doi.org/10.1016/S0010-0277(02)00185-3

Goel, V., Navarrete, G., Noveck, I. A., & Prado, J. (2017). Editorial: The reasoning brain: The interplay between cognitive neuroscience and theories of reasoning. *Frontiers in Human Neuroscience, 10*, Article 673.

Grushcow, M. (2008). *Cognitive Test Software (Version 2.0)* [Computer software]. NTT Systems Inc.

Guo, L., Trueblood, J. S., & Diederich, A. (2017). Thinking fast increases framing effects in risky decision making. *Psychological Science*, *28*(4), 530–543. https://doi.org/10.1177/0956797616689092

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, *24*(12), 2409–2419. https://doi.org/10.1177/0956797613492984

Houdé, O., Pineau, A., Leroux, G., Poirel, N., Perchey, G., Lanoë, C., Lubin, A., Turbelin, M.-R., Rossi, S., Simon, G., Delcroix, N., Lamberton, F., Vigneau, M., Wisniewski, G., Vicet, J.-R., & Mazoyer, B. (2011). Functional magnetic resonance imaging study of piaget's conservation-of-number task in preschool and school-age children: A neo-piagetian approach. *Journal of Experimental Child Psychology*, *110*(3), 332–346. https://doi.org/10.1016/j.jecp.2011.04.008

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience*, *1*(2), 91–100. https://doi.org/10.1038/35039043

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, *23*(1), 315–341. https://doi.org/10.1146/annurev.neuro.23.1.315

Lawson, M. A., Larrick, R. P., & Soll, J. B. (2020). Comparing fast thinking and slow thinking: The relative benefits of interventions, individual differences, and inferential rules. *Judgment and Decision Making*, *15*(5), 660–684.

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain*, *137*(1), 12–32. https://doi.org/10.1093/brain/awt162

Lenartowicz, A., Kalar, D. J., Congdon, E., & Poldrack, R. A. (2010). Towards an ontology of cognitive control. *Topics in Cognitive Science*, *2*(4), 678–692. https://doi.org/10.1111/j.1756-8765.2010.01100.x

Lenhard, W., & Lenhard, A. (2016). *Calculation of effect sizes*. Psychometrica. https://doi.org/10.13140/RG.2.2.17823.92329

Mevel, K., Borst, G., Poirel, N., Simon, G., Orliac, F., Etard, O., Houdé, O., & De Neys, W. (2019). Developmental frontal brain activation differences in overcoming heuristic bias. *Cortex*, *117*, 111–121. https://doi.org/10.1016/j.cortex.2019.03.004

Miller, E. K. (2000). The prefontral cortex and cognitive control. *Nature Reviews Neuroscience*, *1*(1), 59–65. https://doi.org/10.1038/35036228

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Newman, I., 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

Noveck, I. A., Goel, V., & Smith, K. W. (2004). The neural basis of conditional reasoning with arbitrary content. *Cortex*, *40* (4-5), 613–622. https://doi.org/10.1016/S0010-9452(08)70157-6

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2014). Cognitive style and religiosity: The role of conflict detection. *Memory & Cognition*, *42*(1), 1–10. https://doi.org/10.3758/s13421-013-0340-7

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

Prado, J., Chadha, A., & Booth, J. R. (2011). The brain network for deductive reasoning: A quantitative meta-analysis of 28 neuroimaging studies. *Journal of Cognitive Neuroscience*, *23*(11), 3483–3497. https://doi.org/10.1162/jocn_a_00063

Prado, J., & Noveck, I. A. (2007). Overcoming perceptual features in logical reasoning: A parametric fMRI study. *Journal of Cognitive Neuroscience*, *19*(4), 642–657. https://doi.org/10.1162/jocn.2007.19.4.642

Raoelison, M. T., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*, 104381. https://doi.org/10.1016/j.cognition.2020.104381

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004a). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Ridderinkhof, K. R., Van Den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004b). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140. https://doi.org/10.1016/j.bandc.2004.09.016

Rushworth, M. F., Buckley, M. J., Behrens, T. E., Walton, M. E., & Bannerman, D. M. (2007). Functional organization of the medial frontal cortex. *Current Opinion in Neurobiology*, *17*(2), 220–227. https://doi.org/10.1016/j.conb.2007.03.001

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., & Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, *440*(7084), 676–679. https://doi.org/10.1038/nature04513

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shipley-2*. Western Psychological Services.

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*, *24*(4), 423–444. https://doi.org/10.1080/13546783.2018.1459314

Stanovich, K. E., & West, R. F. (1998). Individual differences in framing and conjunction effects. *Thinking & Reasoning*, *4*(4), 289–317. https://doi.org/10.1080/135467898394094

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. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*(4), 672–695. https://doi.org/10.1037/0022-3514.94.4.672

Stollstorff, M., Vartanian, O., & Goel, V. (2012). Levels of conflict in reasoning modulate right lateral prefrontal cortex. *Brain Research*, *1428*, 24–32. https://doi.org/10.1016/j.brainres.2011.05.045

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

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

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

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), 1275–1289. 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

Tsujii, T., Masuda, S., Akiyama, T., & Watanabe, S. (2010). The role of inferior frontal cortex in belief-bias reasoning: An rTMS study. *Neuropsychologia*, *48*(7), 2005–2008. https://doi.org/10.1016/j.neuropsychologia.2010.03.021

Tsujii, T., & Watanabe, S. (2010). Neural correlates of belief-bias reasoning under time pressure: A near-infrared spectroscopy study. *NeuroImage*, *50*(3), 1320–1326. https://doi.org/10.1016/j.neuroimage.2010.01.026

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities.. *Psychological Bulletin*, *133*(6), 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

van Veen, V., & Carter, C. S. (2006). Conflict and cognitive control in the brain. *Current Directions in Psychological Science*, *15*(5), 237–240. https://doi.org/10.1111/j.1467-8721.2006.00443.x

Vartanian, O., Beatty, E. L., Smith, I., Blackler, K., Lam, Q., Forbes, S., & De Neys, W. (2018). The reflective mind: Examining individual differences in susceptibility to base rate neglect with fmri. *Journal of Cognitive Neuroscience*, *30*(7), 1011–1022. https://doi.org/10.1162/jocn_a_01264

Wager, T. D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: A meta-analysis. *NeuroImage*, *22*(4), 1679–1693. https://doi.org/10.1016/j.neuroimage.2004.03.052

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. Harcourt Brace Psychological Corporation.