The temporal dynamics of base rate neglect: No Intuitive Statisticians

Eoin Travers (eoin.travers@sas.ac.uk)
Institute of Philosophy, University of London
Senate House, London WC1E 7HU, UK

Jonathan J. Rolison (jrolison@essex.ac.uk)
Department of Psychology, University of Essex
Wivenhoe Park, Colchester, CO4 3SQ, UK

Aidan Feeney (a.feeney@qub.ac.uk)
School of Psychology, Queen's University Belfast
University Road, Belfast BT7 1NN, UK

Abstract

According to a classic view of reasoning, intuition is fast and effortless but fallible, while reflection is slow and effortful, but reliable. Biases, therefore, emerge when a reasoner's intuitions are systematically wrong and they fail to recognise the need to reflect on them. Recent evidence, however, suggests that people are to some degree aware when their intuitions are wrong, leading to slower responses, and reduced confidence. A possible explanation for this is that our intuitions actually cue multiple conflicting responses, including the correct one, but the wrong one dictates our behaviour unless we consciously detect and resolve this conflict. Here, we tested this explanation using base rate neglect problems, commonly used to assess conflict in reasoning, and recording participants' mouse cursor movements as they chose between possible answers to the problems under time pressure. Descriptions affected both participants' early mouse movements and ultimate responses, and interfered with their use of the base rates, while base rates rarely interfered with participants' use of descriptions, and when they did their influence was at a later point in time. Thus, despite suggestive findings elsewhere, our results support the classic of view reasoning about these problems.

Keywords: Intuition; Reflection; Intuitive logic; Statistical reasoning.

Adults' judgement and reasoning is not as effective as it could be. High-level decision making has been shown again and again to be subject to a range of biases (Evans, Barston, & Pollard, 1983; Gilovich, Griffin, & Kahneman, 2002; Tversky & Kahneman, 1974). A common view is that these biases arise because we often rely on fast, effortless, and automatic Type 1 or *intuitive* processes, which often make use of mental heuristics or rules of thumb, rather than slow, effortful Type 2 or *reflective* processes, which can apply the rules of logic, probability theory, or whatever normative standard is necessary to solve the problem at hand (Evans & Stanovich, 2013; Kahneman & Frederick, 2005).

According to most versions of this account (e.g., Evans, 2006), reasoners rely on intuitive processes by default as they are fast and easy to apply, and produce satisfactory responses in most situations. In contexts where intuitive processes perform poorly reflective processes can often do better, at the cost of greater investment of time and effort. In order to decide when to engage in reflective processing, a lax monitoring hypothesis has been suggested (Evans, 2006; Kahneman & Frederick, 2005). Under this view, people use just enough reflective processing to monitor their intuitive responses for errors, but do so in a lazy way, and often fail to detect situations where they should intervene. Therefore, reasoners are both biased, because they rely on intuitive processes, and unaware of this bias, because they fail to properly monitor those processes.

There is good behavioural evidence, however, that people are at least somewhat aware of when their intuitive responses conflict with normative standards. This evidence comes from studies where classic conflict problems have been compared to no-conflict problems, modified so that the intuitively-appealing response now agrees with the base rate. For instance, consider the lawyer/engineer problem (this version from De Neys, Vartanian, & Goel, 2008):

> In a group, there are 5 engineers, and 995 lawyers.
>
> Jack is a randomly chosen member of this group.
>
> Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles.
>
> Is Jack an engineer, or a lawyer?

Jack's description more closely matches our stereotypes about engineers than those about lawyers (Jack is more *representative* of the "engineer" category), and according to the standard view intuitive processes quickly and automatically indicate this (cf. Barbey & Sloman, 2007). Reflective processes, however, are required to recognize that if 99.5% of the group are lawyers, a randomly chosen member of the group is vastly more likely to be a lawyer then an engineer. As intuitive and reflective processes support different responses here, this is a classic conflict problem. The corresponding no-conflict is created by reversing the base rates, so that 99.5% of the group are engineers. Nevertheless, responses consistent with the description on both problems are generated intuitively. These *base rate problems* are a useful tool for studying conflict in reasoning more broadly.

In studies using conflict and no-conflict versions of this and other reasoning problems, on conflict trials participants are slower to respond (De Neys et al., 2008), are less confident in their responses (De Neys, Cromheeke, & Osman, 2011; De Neys, Rossi, & Houdé, 2013), show an elevated galvanic skin response (De Neys, Moyens, & Vansteenwegen, 2010), and show greater activation in the anterior cingulate cortex, a neural region implicated in the detection of conflict (De Neys et al., 2008), among other subtle indicators of conflict. Thus, it appears that people are, at least under some circumstances, and at least implicitly, aware that their intuitive responses are inadequate.

Less clear, however, is how this awareness comes about. One possible mechanism (De Neys, 2012, 2014; Pennycook, Fugelsang, & Koehler, 2015) is that intuitive processes can partially cue multiple responses simultaneously (a well-documented phenomenon in tasks that do not require reflective engagement at all; see, e.g., Botvinick, Cohen, & Carter, 2004). Thus, if intuitive responses cue only one response, that response is given. If multiple conflicting responses are cued, either a) they will compete until the strongest response is selected, or b) the conflict will be detected, and top-down reflective processes will be engaged to resolve the competition, possibly in favour of the initially weaker but correct response (Pennycook et al., 2015).

This intuitive logic account makes a particularly strong claim: that the correct responses to rather complex reasoning and decision making problems can be generated intuitively and automatically, and the role of top down processes is primarily to inhibit incorrect responses so the correct answer can be given. As a result, the effects described above – indicating greater conflict on conflict problems than no-conflict problems – occur not merely because participants are less certain of their intuitive responses, but because they are actively drawn towards the competing correct response. There is some evidence that this may be the case. Pennycook, Trippas, Handley, & Thompson (2014), for instance, explicitly asked participants to base their responses to base rate problems on either the base rates or on the descriptions, and in both cases found participants were slower to respond, and less confident, when the alternative cue conflicted with their response.

However, the mechanism by which these correct responses are generated is not clear, in particular for complex problems. For instance, in the Cognitive Reflection Task (Frederick, 2005), a set of mathematical problems where the intuitively-given response is incorrect, participants have been found to be slower to respond, and less confident in their responses on the original conflict items than on no-conflict versions where the intuitive response is also correct (De Neys et al., 2013). In Travers, Rolison, and Feeney (2016), however, we presented participants with a four-option multiple choice version of the CRT, and recorded participants' mouse cursor movements as they chose between the competing response options. We found that when participants selected the correct option on conflict problems, they were initially drawn towards the incorrect but intuitively-appealing option before doing so. On the other hand, they were not drawn towards the correct option on trials where they ultimately selected the incorrect, intuitive one, as would be expected if the correct option was partially cued by intuitive processes. Therefore, while participants are to some degree aware that their incorrect intuitive responses may be problematic, it is not clear if intuitive processes actually generate the correct response.

In the current study, we extend this approach to consider the role of intuitive processes in another kind of reasoning problem: base rate neglect problems, such as the lawyer-engineer problem, above. We do so by recording participants' mouse cursor movements as they choose between the response option cued by the description and the option cued by base rate statistics.

According to most accounts, in order to use the base rates on conflict trials participants must inhibit the response initially cued by the description. Therefore, our first goal is to investigate whether this is reflected in participants' cursor movements. Specifically, when base rates do support a response, descriptions can either agree, be neutral, or conflict with them. Early cursor movements should be driven primarily by the contents of these descriptions, while later movements and ultimate responses should be driven either by descriptions in cases where these have not been inhibited, or by base rates otherwise.

However, if the stronger version of the intuitive logic theory is true, and if cursor movements do reflect participants' attraction towards each response option over time, we should also see the reverse pattern: when the descriptions support a response, base rates can either agree, be neutral or conflict, and both early movements and final responses should be driven by both cues, as both are processed intuitively. On the other hand, if the base rate information is not processed intuitively, its influence on participants' actions should emerge later on, and would likely be less than that of the descriptions.

Moreover, De Neys et al. (2008) reported slower response times for description-consistent responses to conflict problems, compared to non-conflict problems, suggesting that base rates were processed, but did not dictate participants' final responses. Here, we will seek to replicate this result, and further to test whether it is driven by participants actually moving initially towards the base rate-cued option on conflict problems, or merely being slower to move towards the description-cued option.

Methods

Participants 50 undergraduate students at Queen's University Belfast took part in the experiment in return for course credit.

Stimuli Reasoning items were adapted from De Neys et al. (2008), and consisted of 40 problems, each made up of base rates of two social categories within a population of 1000 people, and a description of a randomly chosen member of that sample. Participants were asked to use this information to decide which social category the chosen individual was most likely to belong to. The base rate information could either indicate a large majority of the sample belonging to one of the groups (995 vs. 5), or an equal distribution of the two groups (500 vs. 500). The description could either match common stereotypes about one or other group, or be totally uninformative (Table 1).

Questions were presented in four conditions, with 10 questions in each: the base rate and description could both suggest the same response (agree trials), or could disagree (conflict trials), or either the base rate (description only trials) or the description (base rate only trials) could be uninformative. While the same 10 questions, with deliberately uninformative descriptions, were used for the base rate only trials across all conditions, the remaining 30 questions were randomly assigned to the other conditions for each participant.

Table 1: Example reasoning problems from each condition. *Note:* †Category supported by description.

Condition	Base Rates	Description
Agree	†995 Engineers; 5 Lawyers	Jack is 36. He is not married and is somewhat introverted. He likes to spend his free time reading science fiction and writing computer programs.
Conflict	995 violinists; †5 rappers	Jason is 20. He grew up in a poor family in a neglected neighbourhood in Birmingham, and didn't finish his A-levels.
Base rate only	995 Rolling Stones fans; 5 Beatles fans	Mark is 43 years old. He weighs 12 stone and is 5 ft 9 inches tall. He has one younger brother and lives in Manchester.
Description only	†500 60-year-olds; 500 30-year-olds	Gladys is a quiet woman. She lives in a little house with her Yorkshire Terrier where she spends most of her time knitting.

Procedure The 40 trials were presented in a random order for each participant. At the beginning of each trial, the base rate information was presented in the top centre of the screen. After 4 seconds, the description text was also displayed, below the base rates. After a further 5 seconds, a button marked "NEXT" appeared in the bottom centre of the screen. On clicking this, a fixation cross was displayed for 600 milliseconds, after which the probe question was shown in large font in the centre of the screen, the two possible response options appeared as buttons in the top left and right corners, and a timer was displayed below the probe question, which filled up from the bottom over the course of the next 6 seconds, in which time participants were required to give their response by clicking on one of the response buttons. With the onset of the probe question, the mouse cursor was reset to the position of the "NEXT" button if it had left it during the fixation period.

Following trials in which no response was given within 6 seconds, participants were asked to try to respond more quickly. On trials in which participants did not move the mouse cursor from the "NEXT" button within 2 seconds of the onset of the probe question, a message remind them that they were under time pressure, and asking them to "try to start moving as soon as you see the target", was shown, following previous mouse tracking studies (Hehman, Stolier, & Freeman, 2015).

Analysis Pilot studies and previous applications of the mouse tracking paradigm to high level reasoning (i.e. Travers, Rolison, Bright, & Feeney, 2015) indicate that participants' typically move the mouse cursor directly towards one or other response option, but often change direction to select the opposite option, and this proved the case here as well. Therefore, we coded both the initial option participants moved towards and their final choice, on each trial (see Travers, Feeney, & Rolison, in submission, for technical details), and examined the influence of the descriptions and base rates on each. We also performed a more fine-grained analysis of participants' actual movements over time, analysing the proportion of trials where the cursor was on the side of the screen containing each response option, where available, separately for every 20 msec window, and identifying the points from which each cue had a significant influence.

Following previous intuitive logic studies, we also derived a number of measures of conflict or uncertainty, on each trial, including reading time (time spent reading the descriptions before proceeding to the question, beyond the 5 second minimum), response time after the onset of the question, and whether participants' initially moved towards the alternative response. We used these to investigate whether the conflict participants experienced giving the base rate-cued response was affected by the contents of the description, and vice versa. As some participants gave no base rate responses in some conditions, we used mixed effects regression models, with random intercepts for each participant, and each problem.

For brevity, we report unadjusted p values for pairwise comparisons. All significant comparisons remained significant after full Bonferroni adjustment.

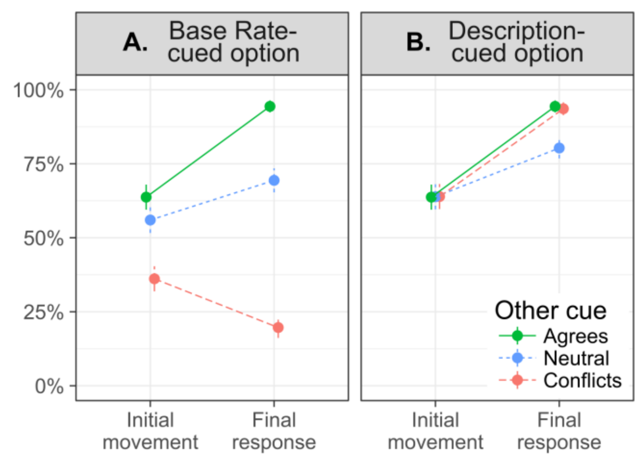

Figure 1. The probability of participants initially moving towards, and ultimately selecting, each response option, as a function of the contents of the other cue. Error bars show SEM.

Results

Figure 1A shows how often participants moved towards or ultimately selected the base rate-cued option when it as available, as a function of the descriptions. Descriptions affected both participants' initial movements, $F(2, 98) = 25.63$, $\eta^2 = 0.26$, $p < .001$, and their final responses, $F(2, 98) = 231.04$, $\eta^2 = 0.71$, $p < .001$, with the effect being greater for final responses, interaction $F(2, 98) = 68.49$, $\eta^2 = 0.20$, $p < .001$. All pairwise comparisons were significant at $p < .001$, except the difference between the effect of agreeing and neutral descriptions on initial movements, $p = .052$. Participants were more likely to finally select the base rate-cued option than to initially move towards it when the descriptions agreed or were neutral, indicating an increased use of the available information over time, but less likely when they disagreed, reflecting a greater increase in the use of the descriptions than in the use of base rates, p's < .001.

Figure 1B, conversely, shows how often participants moved towards or selected the description-cued option as a function of the base rates. Descriptions affected participants' final responses, $F(2, 98) = 13.30$, $\eta^2 = 0.15$, $p < .001$, but not their initial movements, $F(2, 98) < .001$, $\eta^2 < .001$, $p > .99$, BF in favour of null = 15.2; interaction $F(2, 98) = 9.16$, $\eta^2 = 0.03$, $p < .001$. The difference in final responses was driven by the significant differences between conflicting base rates and the other two conditions (neutral and agreeing), p's < .001. Participants were significantly more likely to select the description-cued option than to move towards it in general, and within each condition, p's < .001.

Turning to participants actual movements over time (Figure 2), movements towards the base rate-cued option (2A) were significantly affected by the description from 520 msec, with a significant difference between the conflicting descriptions and the other two conditions from 540 msec, and between neutral descriptions and descriptions that agreed with the base rate from 660 msec. Base rates significantly affected movements towards the description cued option (2B) from 800 msec, driven by the difference between base rates that conflicted with the descriptions and the other two conditions.

Analysis of our measures of conflict showed that when participants did give the base rate-cued response, their reading and response times were significantly affected by the descriptions, χ^2's(2) > 20, p's < .0001, and their initial movements were marginally affected, $\chi^2(2) = 5.28$, $p = .072$. When participants gave the description-cued response, however, their reading and response times, and the initial movements, were not affected by the base rates, $\chi2'$s(2) < 4, p's > .18, or F's(2, 96) < 2, p's > .14, BFs in favour of null hypothesis = [9.46, 9.88, 13.32].

As significant effects of base rates on response times for description-cued responses have been reported elsewhere (e.g. De Neys & Glumicic, 2008; De Neys, Vartanian, & Goel, 2008; Pennycook et al., 2014), we conducted a number of further analyses here. First, we identified participants as either consistently description-driven in their responses

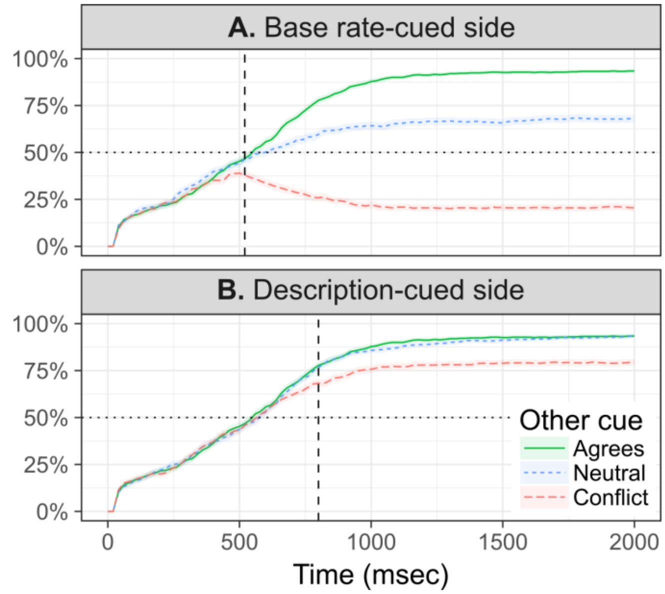

Figure 2. Mouse cursor movements towards the option cued by the base rate **(A)** or by the description **(B)**, as functions of the contents of the other cue. Shading shows SEM.

(>= 9/10 description-cued responses on conflict trials, N = 26), or mixed (N=24; see Figure 3), and repeated the analyses separately for each group. No significant effects were found for either group, BFs = [9.0, 8.0, 6.1] for the consistently description-driven responders, and [3.4, 4.3, 81] for the mixed responders. Second, we combined reading and response times to form a composite measure, but again found no effects.

Finally, we conducted a more in-depth exploratory analysis of the conflict trials, estimating parameters by fitting intercept-only mixed effects models, with random intercepts for each participant and each problem. This approach yields more precise estimates than simply reporting the observed proportions and their standard deviations, although not necessarily producing the same point estimates when individual differences are not symmetrical. When both cues conflicted, there was a 65% probability of initially moving towards the description-cued option, CI = [59%; 71%], and a 90% [82%; 95%] probability of ultimately selecting this option (see also Figure 3). Participants who initially moved towards the description-cued option had a 9% [4%; 15%] probability of changing direction to select the base rate-cued option instead, while those who initially moved towards the base rate-cued option had an 80% [67%; 92%] probability of changing direction to select the description-cued option.

Discussion

We found that descriptions had a robust and rapid influence on both participants' initial cursor movements and their final responses, and overwhelmed the base rate information when they contradicted it. However, base rates had no influence on participants' early movements, only a minor effect on some final responses, and apparently no

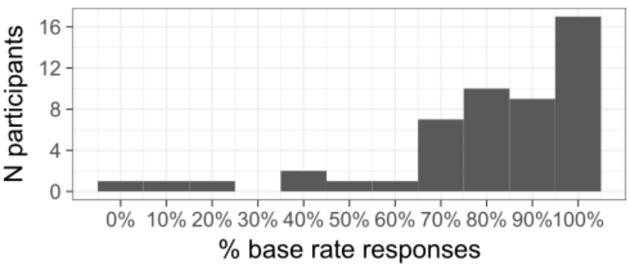

Figure 3: Individual differences in the number of base rate-consistent responses given on conflict trials. Most participants gave base rate-consistent responses on most or all conflict trials, with a small number regularly giving the base rate-consistent response.

effect on response times on trials where they did not dictate the final response.

These results indicate, first of all, that the influence of descriptions on the early stages of reasoning, predicted by standard dual process accounts, is captured by participants' early cursor movements. However, the results go against a strong version of the intuitive logic theory, which would predict that descriptions should influence cursor trajectories in a similar way. Instead, it appears that descriptions are processed quickly, easily, and automatically, and always influence either ultimate responses or subtler measures of conflict. Base rates, however, require slower, effortful, reflective processing, and were only influential on the subset of trials where participants did engage in such processing, and were ignored otherwise. As such, our results are more consistent with a classical dual process account.

These results are somewhat surprising in light of previous work suggesting that participants are slower to give the description-cued response when it conflicts with the base rate (e.g., De Neys & Glumicic, 2008; De Neys et al., 2008; Pennycook et al., 2014). Indeed, our aim here was to explore whether this slowdown was due to participants being initially drawn towards the base rate-cued option, in the same way as they are initially drawn towards the description-cued option, or if they merely slowed their responses without making overt base rate-driven movements. Although space precludes a full meta-analysis here, it is worth noting that this experiment differed from previous intuitive logic studies using base rate neglect problems in a number of ways. For instance, we required participants to respond under moderate time pressure – six seconds to respond, and two to initiate a movement – and most responses took approximately two seconds, markedly less than in previous work using these problems. This may have reduced participants' use of the base rates, although this would perhaps be surprising under the strong intuitive logic account, which holds that base rates can be processed quickly and automatically as well. In fact, only two other studies have asked participants to solve these problems under time pressure. De Neys et al (2008), in an fMRI study, found slower response times for these problems, but tested only 13 participants, and found effects that were just on the threshold of statistical significance ($p \approx .045$, based on reported $p_{rep} = .99$). Pennycook et al. (2014; Experiment 2) also asked participants to respond quickly, and found similar effects, but explicitly told participants to rely on the base rates on some trials and the descriptions on others. As such, it is perhaps unsurprising that their participants were more affected by base rates than ours, who had to spontaneously decide to rely on them.

Another difference was due to the restrictions of the mouse tracking paradigm. Participants saw the base rates first, followed by the descriptions, followed by the questions, in order to isolate the actual decision from initial processing of the background information. It is possible that participants' did experience some conflict, but that this occurred before the onset of the questions, and was not captured by our analysis of participants' reading times. Pilot studies, however, showed that participants' mouse cursor trajectories were uninformative when problems were presented all at once.

More generally, we found that on conflict problems our participants gave the base rate-cued response only 20% of the time, less than in previous work. It is plausible that our participants used base rates less overall, perhaps accounting for their failure to influence response times. However, we did take a number of steps based on previous work to encourage participants to make use of the base rate information, including presenting the base rates in a frequency format (995/1000) rather than as percentages (99.5%), using extreme base rate frequencies, and emphasizing the random selection of the person from the group. In this respect, we gave our participants every opportunity we could to be influenced by the base rates. Elsewhere, however, individual differences have previously been observed in both participants' use of base rates in general (Stanovich & West, 2000) and in the strength of intuitive logic effects on this and other tasks (De Neys & Bonnefon, 2013). It is therefore possible that our cohort of participants, for whatever reason, was less inclined to make use of base rates in reasoning.

Finally, it is also worth noting that our findings are not inconsistent with single-process accounts of reasoning, where there is no qualitative difference between intuition and reflection (e.g. Kruglanski & Gigerenzer, 2011). However, descriptions influence the reasoning process earlier than base rates do, and appear to play a role in all decisions, whereas base rates only influenced a minority of trials. As such, these results, as well as other findings using this paradigm, place certain constraints on single-process theories of reasoning. Ultimately, we believe that the actual cognitive and neural underpinning of human reasoning are considerably more complex than current accounts, whether they rely on one, two, or many kinds of processing. However, the distinction between intuitive and reflective thinking is a useful one, and so outlining the capabilities and limitations of intuition remains an important task. In this work, we have confirmed that processing descriptions is well within those capabilities, but, contrary to some previous work, we found evidence that drawing inferences from base rate statistics may not be.

References

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*(03). https://doi.org/10.1017/S0140525X07001653

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

De Neys, W. (2012). Bias and Conflict: A Case for Logical Intuitions. *Perspectives on Psychological Science*, *7*(1), 28–38. https://doi.org/10.1177/1745691611429354

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

De Neys, W., & Bonnefon, J.-F. (2013). The 'whys' and 'whens' of individual differences in thinking biases. *Trends in Cognitive Sciences*, *17*(4), 172–178. https://doi.org/10.1016/j.tics.2013.02.001

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in Doubt: Conflict and Decision Confidence. *PLoS ONE*, *6*(1), e15954. https://doi.org/10.1371/journal.pone.0015954

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., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: autonomic arousal and reasoning conflict. *Cognitive, Affective & Behavioral Neuroscience*, *10*(2), 208–216. https://doi.org/10.3758/CABN.10.2.208

De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, *20*(2), 269–273. https://doi.org/10.3758/s13423-013-0384-5

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.

Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, *13*(3), 378–395. https://doi.org/10.3758/BF03193858

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

Evans, J. S. B. T., & 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

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

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning*. Cambridge University Press.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*(1), 97–109. https://doi.org/10.1037/a0020762

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

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 544–554. https://doi.org/10.1037/a0034887

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

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition*, *150*, 109–118. https://doi.org/10.1016/j.cognition.2016.01.015

Travers, E., Feeney, A., & Rolison, J. J. (in submission). Similarity, and then Conceptual Knowledge drive Inductive Reasoning.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124