



Brief article

Are we good at detecting conflict during reasoning?

Gordon Pennycook*, Jonathan A. Fugelsang, Derek J. Koehler

Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1

ARTICLE INFO

Article history:

Received 12 September 2011

Revised 12 April 2012

Accepted 12 April 2012

Available online 8 May 2012

Keywords:

Base-rate neglect
 Dual-process theory
 Conflict detection
 Heuristics
 Biases
 Reasoning

ABSTRACT

Recent evidence suggests that people are highly efficient at detecting conflicting outputs produced by competing intuitive and analytic reasoning processes. Specifically, De Neys and Glumicic (2008) demonstrated that participants reason longer about problems that are characterized by conflict (as opposed to agreement) between stereotypical personality descriptions and base-rate probabilities of group membership. However, this finding comes from problems involving probabilities much more extreme than those used in traditional studies of base-rate neglect. To test the degree to which these findings depend on such extreme probabilities, we varied base-rate probabilities over five experiments and compared participants' response time for conflict problems with non-conflict problems. Longer response times for stereotypical responses to conflict versus non-conflict problems were found only in the presence of extreme probabilities. Our results suggest that humans may not be consistently efficient at detecting conflicts during reasoning.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

According to dual-process theories (Evans, 2008; Frankish & Evans, 2009; Stanovich, 2004), reasoning and decision-making are based on two qualitatively different types of cognitive processes (Evans, 2009, 2011a): Type 1 (T1) processes are fast, intuitive, and high capacity whereas Type 2 (T2) processes are slow, deliberative, and analytical. They contrast, in that, T1 processing provides quick default outputs while T2 processing is engaged deliberately to manipulate explicit representations via working memory (Evans, 2010).

A key finding is that the output of T1 processing often dominates reasoning such that participants frequently exhibit poor performance on problems that elicit an incorrect intuitive response (Evans, 2008). This performance deficit is usually explained by a lack of T2 override – i.e., participants are unable to engage in sufficient analytic processing to overcome their initial intuitive response (Evans, 2008; Frankish & Evans, 2009; Kahneman, 2003; Stanovich, 2004).

1.1. Dual-process theory and base-rate neglect

A classic finding from the heuristics and biases research program is base-rate neglect (Kahneman & Tversky, 1973). Consider the following problem (from De Neys and Glumicic (2008)):

In a study 1000 people were tested. Among the participants there were 995 nurses and 5 doctors. Paul is a randomly chosen participant of this study.

Paul is 34 years old. He lives in a beautiful home in a posh suburb. He is well spoken and very interested in politics. He invests a lot of time in his career.

What is most likely?

- (a) Paul is a nurse.
- (b) Paul is a doctor.

For this problem, the personality description suggests a different response (“doctor”) than does the base-rate information (“nurse”). Many studies have shown that people, when faced with such problems, tend to base their judgments on the individuating information (here, the personality description) and neglect or underweight the base-rate (for a review, see Barbey & Sloman, 2007). The question addressed here is whether people recognize the conflict

* Corresponding author.

E-mail address: gpennyco@uwaterloo.ca (G. Pennycook).

between the individuating and base-rate information. Note that since the individuating information in the above problem does not have a numerical value, there is no actual correct response. Bayes theorem requires the base-rate probability to be integrated with the individuating information for a normative response. Our primary concern is not whether participants are accurate in solving these problems, but whether they will detect the inherent conflict. For this reason, responses will be categorized as “base-rate” or “stereotypical” answers as opposed to correct or incorrect. This also allows us to avoid the confusion that sometimes occurs in the literature where a normatively correct response is assumed to have arisen out of T2 processing and a biased (stereotypical in this case) response is assumed to have arisen out of T1 processing (see Evans, 2011b).

De Neys and colleagues have argued that people are aware, albeit implicitly, when their reasoning processes are subject to biases, such as base-rate neglect (De Neys, Comheeke, & Osman, 2011; De Neys & Franssens, 2009; De Neys & Glumicic, 2008; De Neys, Vartanian, & Goel, 2008). This argument has important implications for reasoning because it suggests that such reasoning biases are caused primarily by a failure of inhibition and override rather than a failure of conflict detection (e.g., an absence of T2 processing). The first and most basic piece of evidence with respect to base-rate neglect came from response time data. De Neys and Glumicic (2008) showed that participants took longer to answer base-rate problems when the personality description and base-rates were incongruent (in conflict) than when they were congruent (non-conflicting). The crucial assumption here is that increased response time reflects slow T2 processing. Thus, participants apparently detected the conflict and engaged in increased T2 processing to resolve it. Strikingly, response times were longer for incongruent than congruent trials even when the analysis is limited to trials in which the stereotypical or intuitive response was chosen, which was taken to indicate that the conflict with the base-rate was detected but not overridden in producing a response. This finding was reinforced by additional measures (e.g., memory retrieval, verbal protocol analyses), leading De Neys and Glumicic to conclude that, “at least in case of the classic base rate neglect phenomenon, heuristic thinking seems to be always accompanied by successful analytic monitoring” (p. 1280).

There is, however, reason to question the generalizability of these findings. Specifically, the base-rates used by De Neys and colleagues are substantially more extreme (995/996/997 out of 1000) than those used in past base-rate research (e.g., 70 out of 100 in Kahneman and Tversky's (1973) lawyer–engineer problem). In other base-rate problems (such as the Tom W. problem; Evans, Handley, Over, & Perham, 2002; Kahneman & Tversky, 1973), explicit numerical base-rates are not presented at all; rather, participants must estimate them based on their own general knowledge. While use of extreme base-rates may help to avoid complications with respect to the normative status of the base-rate response (as discussed above; see also Franssens & De Neys, 2009), an important question remains: How obvious does a conflict between a base-rate

and individuating information need to be for it to be detected? If conflict detection is highly efficient, participants should be able to detect a conflict between a stereotypical personality description and an implicit or moderate base-rate probability, not just an extreme one.

To address this question, we created a set of stimuli concerning university academic majors that varied in terms of students' beliefs about their relative program size (i.e., base-rate) and stereotypical characteristics of the prototypical student. The base-rates were then either left implicit (Experiment 1), made explicit at a moderate level (e.g., 70%/30%; Experiment 2) or made explicit at an extreme level (e.g., 995/5; Experiment 3). Additionally, we replicated our findings using De Neys and Glumicic's (2008) original stimuli with extreme base-rates (e.g., 995/5; Experiment 4) and moderate base-rates (e.g., 70/30; Experiment 5). If conflict detection is highly efficient across a broad range of base-rate problems, response times when choosing the stereotypical option should be longer when it has a low base-rate probability (incongruent problems) than when it has a high base-rate probability (congruent problems) across each experiment. Alternatively, if conflict detection occurs only in the presence of factors that draw attention to base-rates (i.e., explicit presentation of extreme base-rates), then such longer response times may not hold consistently across all experiments.

As noted above, De Neys and colleagues have used other measures in addition to response time in their studies of base-rate neglect. We chose to focus on response time as it seemed to be the most direct and commonly used means of assessing engagement of T2 processing.

2. Method

2.1. Materials

Participants in each experiment completed 18 base-rate problems, in which they predicted which of two possible groups (that differed in base-rate) an individual belonged to. The personality descriptions for each problem consisted of individuating information (stereotypes) that always clearly favoured one group over the other. For congruent problems, the alternative favoured by the stereotype had the higher base-rate. For incongruent problems, the alternative favoured by the stereotype had the lower base-rate. Problems for Experiments 1–3 were constructed using an online pretest¹ and validated behaviorally by participants following Experiment 1.² Personality descriptions for Experiments 1–3 consisted of a list of 5 personality traits (see Supplementary materials).

¹ One-hundred thirty University of Waterloo students ranked (based on enrolment size) and assigned personality traits to various majors on campus.

² Following the experimental trials, participants in Experiment 1 were asked to assign an estimated enrolment proportion for each major of the presented pairs. The mean proportion across all comparisons was 70.2 for the actual larger major. Participants also rated the relative similarity of each personality description to the two majors in each pair. The mean similarity rating across all comparisons was 5.2 out of 7, ($t(51) = 16.98$, $p < .001$ vs. scale midpoint of 4), indicating that the personality descriptions provided diagnostic support of the targeted major in each pair.

2.2. Procedure

In each of the five experiments, problems were presented in full, one at a time in a random order on a computer using E-Prime 1.1 (Schneider, Eschman, & Zuccolotto, 2001). Response time was measured from the onset of problem presentation, and therefore included reading time. Instructions were adapted from De Neys and Glumicic (2008).

2.2.1. Experiment 1: Implicit base-rates

Fifty-two University of Waterloo students (39 female; $m = 19.8$ years) participated for course credit. Numerical base-rates were not explicitly provided, but post-task ratings verified that one of the options in each pair was consistently perceived to have a higher base-rate than the other. Participants answered 18 problems (9 congruent, 9 incongruent).

2.2.2. Experiment 2: Moderate base-rates

Fifty-two University of Waterloo students (41 female; $m = 19.3$ years) participated for course credit. The explicitly-presented base-rates were based on the mean estimated base-rate proportions provided by participants in Experiment 1. Across the 18 problems, the mean base-rate was 70% for the larger major and 30% for the smaller major, ranging by problem from 62%/23% to 82%/18% (see Supplementary materials). Participants answered 18 problems (9 congruent, 9 incongruent).

2.2.3. Experiment 3: Extreme base-rates

Thirty-two University of Waterloo students (23 female; $m = 19.3$ years) participated for course credit. Following De Neys and Glumicic (2008), three base-rate ratios were presented an equal number of times: 995/5, 996/4, 997/3. Participants answered 18 problems (9 congruent, 9 incongruent).

2.2.4. Experiment 4: Replication with extreme base-rates

Thirty-two University of Waterloo students (18 female; $m = 20.5$ years) participated for course credit. Participants completed the 18 base-rate problems (6 congruent, 6 incongruent, 6 neutral) used by De Neys and Glumicic (2008). The neutral problems were constructed by De Neys and Glumicic with personality descriptions that did not contain any stereotypes. For consistency among the experiments, the neutral items were excluded from analysis. Three base-rate ratios were used an equal number of times: 995/5, 996/4, 997/3.

2.2.5. Experiment 5: Replication with moderate base-rates

Forty University of Waterloo students (17 female; $m = 19.9$ years) participated for course credit. Participants completed the same 18 base-rate problems from Experiment 4, this time with three moderate base-rates, 70/30, 71/29, 72/28, each used an equal number of times.

Table 1

Mean choice proportion for high base-rate alternative as a function of problem type for Experiments 1–5.

	Choice proportion	
	Congruent	Incongruent
E1: Implicit	.84	.21
E2: Moderate	.90	.39
E3: Extreme	.95	.59
E4: Extreme	.91	.24
E5: Moderate	.95	.26

3. Results

3.1. Analysis strategy and design

Following De Neys and Glumicic (2008), only response times (RTs) for responses consistent with the base-rate and stereotype for congruent problems were included in the analysis. RTs were analysed using a repeated measures ANOVA with three levels (responses to congruent problems consistent with base-rate and stereotype, responses to incongruent problems consistent with base-rate, responses to incongruent problems consistent with stereotype). RTs were converted to \log^{10} prior to analysis. Only extreme outlying RTs (>4 SDs) were excluded from analysis, representing 1.1% of the data. Some participants were not included in the ANOVA because they gave only base-rate responses or only stereotypical responses for incongruent problems.³

3.2. Choice proportion for high base-rate alternative

The proportion of choices that were consistent with the base-rate for incongruent problems was compared to congruent with paired-samples *t*-tests. A substantial decrease in choice proportion was observed for incongruent problems relative to congruent problems in every experiment, all t 's > 8.25 , all p 's $< .001$ (see Table 1).

3.3. Response time

RT differed as a function of congruency in Experiments 1–4, all F 's > 4.04 , all p 's $< .026$. While the effect was marginal for Experiment 5, $F(2,34) = 2.76$, $MSE = .004$, $p = .078$, the pattern was consistent throughout (see Table 2). There was a general increase in RT for incongruent relative to congruent problems. One reason for this overall main effect is that RTs were longer, in every experiment (all t 's > 2.36 , all p 's $< .028$; see rightmost column of Table 2), when participants faced with incongruent problems chose the high base-rate alternative despite conflicting stereotype information. This is unsurprising, as in such cases participants would need to engage in increased T2 processing to override the intuitive stereotypical response

³ Across the five experiments, 9.6% of participants gave only base-rate responses and 19.7% gave only stereotypical responses. All possible participants are included in the following ANOVAs and *t*-tests (i.e., 9.6% excluded for incongruent-stereotypical and 19.7% excluded for incongruent-base-rate).

Table 2

Mean response time (in seconds) as a function of problem type and response (either consistent with stereotype or base-rate) for Experiments 1–5. Difference (in seconds) between incongruent (stereotypical or base-rate response) and congruent response times are in brackets.

	Congruent	Incongruent	
		Stereotypical (diff)	Base-rate (diff)
E1: Implicit	11.5	11.5 (0.0)	14.7 (3.2 [*])
E2: Moderate	13.9	14.6 (0.7)	17.0 (3.1 [*])
E3: Extreme	11.5	13.7 (2.2 [*])	14.5 (3.0 [*])
E4: Extreme	17.4	19.7 (2.3 [*])	22.5 (5.1 [*])
E5: Moderate	16.0	16.5 (0.5)	18.7 (2.7 [*])

^{*} Represents significance at a .05 level for the incongruent–congruent response time comparison.

(Evans, 2010). The “crucial question”, in the words of De Neys and Glumicic, is whether, compared to congruent problems, RTs were higher for incongruent problems for which participants gave the stereotypical response ($p = .1252$). For this analysis, individual t -tests were used.

3.3.1. Experiment 1: Implicit base-rates

A t -test revealed no significant difference between congruent and stereotype-consistent incongruent RTs, $t(50) = 0.25$, $SE = .016$, $p = .80$. Thus, as Table 2 shows, the overall increase in RT for incongruent problems was caused by an increase for responses consistent with the base-rate information and not by an increase for stereotypical responses. This finding suggests that participants did not reliably detect the conflict between the implicit base-rate information and the stereotypical personality description.

3.3.2. Experiment 2: Moderate base-rates

As in Experiment 1, the crucial difference in RT between congruent problems and stereotype-consistent responses to incongruent problems did not approach significance, $t(47) = 0.43$, $SE = .019$, $p = .667$, suggesting that participants did not reliably detect the conflict between base-rate and stereotypical information, even when the base-rates were explicitly presented. Importantly, this finding cannot be attributed to a total neglect of base-rates because participants actually had a higher proportion of base-rate consistent responses for incongruent problems (39%) than did De Neys and Glumicic’s participants (19–22%).

3.3.3. Experiment 3: Extreme base-rates

Contrary to Experiments 1 and 2, when extreme base-rates were presented there was an increase in RT for both base-rate-consistent responses, $t(29) = 3.59$, $SE = .024$, $p = .001$, and stereotype-consistent responses, $t(22) = 2.46$, $SE = .021$, $p = .022$, relative to congruent problems. Thus, we replicate the key finding of De Neys and Glumicic when extreme base-rates were adopted using our stimuli.

3.3.4. Experiment 4: Replication with extreme base-rates

Again, there was an increase in RT for both base-rate-consistent responses, $t(21) = 2.36$, $SE = .035$, $p = .028$, and stereotype-consistent responses, $t(29) = 2.48$, $SE = .015$, $p = .019$, relative to congruent problems. This replicates

De Neys and Glumicic using their stimuli with extreme base-rates.

3.3.5. Experiment 5: Replication with moderate base-rates

Finally, using De Neys and Glumicic’s stimuli but with moderate base-rates, the crucial difference in RT between stereotype-consistent responses to incongruent problems and congruent problems did not approach significance, $t(35) = 0.67$, $SE = .013$, $p = .50$, again suggesting a general failure of conflict detection in the absence of extreme base-rates. As such, the finding that conflict detection is limited to the case of explicitly-presented extreme base-rates cannot be attributed to the stimuli we developed for Experiments 1–3.⁴

4. General discussion

Previous research on conflict detection during reasoning has indicated that choices exhibiting base-rate neglect are reliably accompanied by detection of a conflict between the available individuating information and the base-rate (De Neys & Glumicic, 2008; De Neys et al., 2008, 2011). This past research, however, focused exclusively on problems with extreme base-rates. The experiments reported here suggest that base-rate probabilities that are implicit (Experiment 1) or moderate (e.g., 70/30; Experiments 2 and 5) do not reliably elicit increased response time when they are “neglected” (i.e., when counter-base-rate choices are made). The apparent lack of reliable conflict detection for all but extreme base-rates suggests that humans are not highly efficient at detecting conflict during reasoning, at least as it applies to base-rate neglect.

In a follow-up, De Neys and Franssens (2009) found evidence for conflict detection using a base-rate of 90 out of 100. While this is less extreme than 995 out of 1000, this difference is fairly small on the probability scale (i.e., 9.5%). The collective results to date suggest that the threshold required for reliable conflict detection during base-rate neglect is some number higher than 70% and lower than 90%. More research is required to further isolate the boundary conditions of this effect.

It is important to emphasize that the present research focused solely on one measure (response time) and one task (base-rate neglect). Researchers have offered converging evidence for conflict detection using tasks other than base-rate problems (for a review, see De Neys, 2012), such as syllogisms (De Neys & Franssens, 2009; De Neys, Moyens, & Vansteenwegen, 2010), the conjunction fallacy (De Neys et al., 2011; Villejoubert, 2009), and ratio bias (Bonner & Newell, 2010). Clearly, there are situations

⁴ As a secondary analysis of the across-experiment pattern, a difference score between congruent and stereotype-consistent incongruent was computed for each experiment (see Table 2). Data for experiments with moderate base-rates (E2 and E5) were merged, as were those with extreme base-rates (E3 and E4). A between-subjects one-way ANOVA was run with three base-rate levels (implicit, moderate, extreme). The effect of base-rate level was significant for raw RT, $F(2, 185) = 3.52$, $p = .032$, and marginally significant for \log^{10} RT: $F(2, 185) = 2.47$, $p = .087$.

where conflict is reliably detected during reasoning. That basic point is not in question here.

Moreover, in their studies of base-rate neglect, De Neys and colleagues have supported their claims using a host of additional measures not used in the present research, such as an increase in memory retrieval for base-rates found in incongruent relative to congruent problems (De Neys & Glumicic, 2008), an increased tendency to review the base-rate information for incongruent relative to congruent problems (De Neys & Glumicic, 2008), a failure to mention the extreme base-rate probabilities during verbal protocol (De Neys & Glumicic, 2008), and a decrease in confidence for incongruent relative to congruent problems (De Neys et al., 2011). In addition, De Neys et al. (2008) used fMRI to show that there was increased activation in the anterior cingulate cortex, which is often referred to as the conflict detection center of the brain (Carter et al., 1998), when participants responded to incongruent relative to congruent problems. While these studies also used extreme base rates, such measures may be more sensitive than response time and therefore could be useful in the exploration of moderate base rates.

Our response-time results do suggest, however, that any characterization of conflict (or “analytic”) detection as highly efficient during base-rate neglect may be overstated in that it implies a level of generalizability over different types of base-rate problems that is not, thus far, empirically supported. In fact, conflict detection in the absence of correction during base-rate neglect may occur only under somewhat limited circumstances. By contrast, reasoning conflicts in everyday life are frequently less obvious even than those represented by the moderate base-rate problems used here. In terms of everyday probabilistic reasoning, base-rate problems where the prior probabilities are not explicitly presented have been argued to be the most commonly encountered in everyday life (Gigerenzer, 2007). Such problems revealed no evidence of conflict detection whatsoever.

Why do only extreme base-rates apparently trigger reliable conflict detection, at least as measured by response time? Recently, De Neys (2012) theorized that conflict detection occurs in situations where people have “logical intuitions”. Specifically, information that has previously been considered to be more analytic in nature, such as base-rate probabilities, may be accessible to intuitive or T1 processing. Under this explanation, the conflict that is detected during base-rate neglect is between two T1 outputs as opposed to a T1–T2 conflict as originally thought. Indeed, Pennycook and Thompson (2012) provide evidence that base-rates are accessible to T1 processing by showing that base-rates influence probability estimates even when participants are asked to give the first response that comes to mind and when there is no conflict between base-rates and individuating information. They also show that using base-rates for intuitive judgments does not increase response time as one would expect if base-rate use requires T2 processing. Taken alongside the current work, this research supports De Neys’ hypothesis that the conflict detected between extreme base-rates and individuating information is really a conflict between two T1 outputs. Under this explanation, moderate base-rates may not draw

enough attention to receive T1 processing at all, or alternatively, may receive T1 processing but not be extreme enough to trigger the high-base-rate response as output of that processing. This suggests that a potentially productive research agenda may entail using basic principles of attention and perception to better understand problem features that lead people to quickly endorse their initial intuitions or to instead subject them to further scrutiny.

While the present work is concerned specifically with base-rate neglect, it may have implications for conflict detection during reasoning more broadly. As mentioned above, researchers have also demonstrated efficient conflict detection using syllogisms, the conjunction fallacy, and ratio bias (De Neys, 2012). If conflict detection during base-rate neglect only occurs when the base-rates are extreme enough to generate an intuitive output, it may be the case that these other tasks also elicit a conflict between intuitive responses. Consistent with this, Morsanyi and Handley (2011) found that people have an intuitive sense about whether syllogisms are valid, which suggests that syllogisms where believability and validity conflict may elicit increased T2 processing because of a conflict between two T1 outputs. Future research is required to address this issue further and to determine if the purported “analytic responses” for ratio bias and the conjunction fallacy actually require T2 processing.

Acknowledgements

We would like to thank Colin MacLeod and Derek Besner for their comments on an earlier version of the manuscript. Funding for this study was provided by the Natural Sciences and Engineering Research Council of Canada.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.cognition.2012.04.004>.

References

- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioural and Brain Sciences*, *30*, 241–256.
- Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition*, *38*, 186–196.
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and online monitoring performance. *Science*, *280*, 747–749.
- De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*, 28–38.
- De Neys, W., Comheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PLoS One*, *6*, e15954.
- De Neys, W., & Franssens, S. (2009). Belief inhibition during thinking: Not always winning but at least taking part. *Cognition*, *113*, 45–61.
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1284–1299.
- De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we’re biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *10*, 208–216.
- De Neys, W., Vartanian, O., & Goel, V. (2008). When our brains detect that we are biased. *Psychological Science*, *19*, 483–489.
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.

- Evans, J. St. B. T. (2010). Intuition and reasoning: A dual-process perspective. *Psychological Inquiry*, 21, 313–326.
- Evans, J. St. B. T. (2011a). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, 31, 86–102.
- Evans, J. St. B. T. (2009). How many dual process theories do we need: One, two, or many? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33–54). Oxford: Oxford University Press.
- Evans, J. St. B. T., Handley, S. J., Over, D. E., & Perham, N. (2002). Background beliefs in Bayesian inference. *Memory & Cognition*, 30, 179–190.
- Evans, J. St. B. T. (2011b). Dual-process theories of reasoning: Facts and fallacies. In K. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning*. New York: Oxford University Press.
- Frankish, K., & Evans, J. St. B. T. (2009). The duality of mind: An historical perspective. In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 1–32). Oxford: Oxford University Press.
- Franssens, S., & De Neys, W. (2009). The effortless nature of conflict detection during thinking. *Thinking and Reasoning*, 15, 105–128.
- Gigerenzer, G. (2007). *Gut feelings*. London, UK: Penguin.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Morsanyi, K., & Handley, S. J. (2011). Logic feels so food—I like it! Evidence for intuitive detection of logicity in syllogistic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <http://dx.doi.org/10.1037/a0026099>.
- Pennycook, G., & Thompson, V. A. (2012). Reasoning with base-rates is routine, relatively effortless and context-dependent. *Psychonomic Bulletin & Review*. <http://dx.doi.org/10.3758/s13423-012-0249-3>.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc..
- Stanovich, K. E. (2004). *The Robot's rebellion: Finding meaning in the age of Darwin*. Chicago: The University of Chicago Press.
- Villejoubert, G. (2009). Are representativeness judgments automatic and rapid? The effect of time pressure on the conjunction fallacy. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 30, 2980–2985.