Title
Instruction type and believability influence on metareasoning in a base rate task

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

Authors
Valerjev, Pavle
Dujmović, Marin

Publication Date
2017

Peer reviewed

Instruction type and believability influence on metareasoning in a base rate task

Pavle Valerjev (valerjev@unizd.hr)
Department of Psychology, University of Zadar, Obala kralja Petra Krešimira IV 2
Zadar, 23000 Croatia

Marin Dujmović (marin.dujmovic@gmail.com)
Department of Psychology, University of Zadar, Obala kralja Petra Krešimira IV 2
Zadar, 23000 Croatia

Abstract

Task dependent conflict has been shown to reduce metacognitive judgements of confidence and prolong response times in various reasoning tasks. For this study a modified version of the base rate task was used to induce conflict while measuring response times and judgements of confidence. The aim of this experiment was to determine the influence of different instruction conditions (reasoning according to belief or according to mathematical probability) on fluency and metacognitive judgements. As expected, participants experienced higher levels of conflict when reasoning according to mathematical probability even though conflict effects were present in both conditions. Additionally, higher believability items mitigated conflict influence while reasoning in accordance with belief and increased it when reasoning in accordance with mathematical probability. These results enrich the growing field of metareasoning research and are discussed as such.

Keywords: metacognition; metareasoning; base rate neglect; conflict monitoring; judgement of confidence

Introduction

The dual process approach in the field of reasoning is based on the hypothesis of two distinct types of processing. Type 1 processing is heuristic based, fast and comes at a low cognitive resource cost, while Type 2 processing is analytic based, slower, less intuitive, and comes at a high resource cost (Evans, 2007). However, this is just a simplified representation of what is a large set of theories. Since this approach encompasses a large number of theories and models we recommend an excellent review by Evans (2012) for a better understanding of all the complexities in this field. Within this framework a cognitive bias can be defined as a dominant Type 1 response when it is not appropriate and results in a normatively incorrect answer. One of the most commonly studied biases is the belief bias in which a believable response is dominant and more acceptable regardless of correctness (Evans, Barston, & Pollard, 1983). Tasks which elicit the belief bias have mostly been used in studies of formal types of reasoning such as syllogistic logical reasoning. However, there are other tasks in which participants are led to reason according to their belief. One such task is a modified version of Kahneman and Tversky's (1973) classic base rate task.

The modified base rate task is interesting to researchers because it is suitable for introducing conflict between processes that give rise to different responses. To better understand this task, examine a simple example.

> Person X is popular.
> Person X is chosen at random from a group consisting of 875 postmen and 125 actors.
>
> The question: Which is more probable?
> 1. Person X is a postman
> 2. Person X is an actor

In the example above there are two sources of information on which a person can base his or her response. The mathematical ratio (the base rate) of postmen to actors would indicate that the first answer is more probable. However, common belief based on personal, everyday experience would indicate actors are more likely to be popular than postmen. This strong association between the trait of popularity and the group of actors leads most participants to choose answer number two (e.g. Obrecht & Chesney, 2016; Pennycook, Fugelsang & Koehler, 2015). By changing the base rate, congruent versions of this task can be constructed and studied in comparison to conflict ones.

Our research is based on the model proposed by Pennycook and his colleagues (2015). They propose that multiple processes generate initial responses to a particular task or problem (Type 1 processes). The initial responses may be congruent or in conflict. Conflict monitoring then detects (or fails to detect) incongruent results of these processes which leads to Type 2 processing (or simple Type 1 processing if no conflict is detected). Type 2 processing then includes both rationalization of the dominant initial response, and what is usually considered pure Type 2 processing – cognitive decoupling (choosing an alternative response). It is important to note that rationalization and decoupling are not necessarily processed on a conscious level. The dominant response is the one with the highest *weight*, which would explain the difference in influence and persistence of those responses even when instructed to reason by different criteria. In the example given above the belief-based response would represent the dominant initial response, while the mathematical probability response would be the second initial response. This model provides clear hypotheses on the influence of conflict on response times, levels of confidence, required cognitive load and others.

As indicated, conflict monitoring and detection are key processes which link Type 1 and Type 2 processes according to the proposed model. For the purposes of this study inducing conflict was of the greatest importance. Conflict monitoring and detection have been identified as important processes in a variety of reasoning tasks (De Neys, 2014; De Neys & Glumicic, 2008). Conflict in base rate tasks is detected even when participants choose the stereotypical answer because it can be observed in prolonged response times (Pennycook et al., 2015). Since conflict detection in these tasks is present regardless of participant response it is interesting how this process effects other psychological constructs and processes.

Metacognition is broadly concerned with knowledge about, the monitoring and evaluation of other mental processes (Nelson & Narens, 1990). Recent research in the field has been focused on possible sources of metacognitive judgements in various reasoning tasks (Markovits, Thompson, & Brisson, 2015; Thompson, Evans, & Campbell, 2013; Thompson & Johnson, 2014; Thompson et al., 2013). Building on older research concerning metamemory Ackerman and Thompson (2015) lay out a framework of metareasoning. They outline a number of different metareasoning judgements. For example, the judgement of solvability refers to the probability a particular task is solvable based on type of task, prior knowledge and experience of the participant. Of the basic metareasoning judgments we focus on the final judgements of confidence. These judgements represent retrospective confidence that a final solution to a problem or task is correct. Typically, metareasoning studies make use of a variety of tasks which include the possibility of both heuristic and analytical responses. The interpretation of results in these studies then naturally takes into account the dual-process approach to reasoning. Participants usually express more confidence for heuristic-based responses (Thompson, Evans, & Campbell, 2013). It has also been established that in these types of tasks participants tend to be overconfident and their judgements are not dependent on accuracy (Thompson et al., 2013). Therefore, metacognitive judgements in these tasks are formed based on other cues such as fluency (the speed and ease of generating responses) and the presence of conflicting answers (Thompson et al., 2013). Conflict in base rate tasks has been shown to decrease final judgements of confidence (Pennycook Trippas, Handley, & Thompson 2014) and prolong response times (Pennycook et al., 2015).

Manipulating the content of tasks is only one way to influence reasoning processes. One of the more interesting manipulations is varying instruction types, giving more weight to a particular type of reasoning or to specific content. Explicit instruction to reason according to logic increases accuracy (Evans, Handley, Neilens, & Over, 2010) as well as confidence ratings (Trippas, Thompson, & Handley, 2016), depending on task difficulty. Within the described model of dual processing, instruction type modifies the weight of initially generated responses therefore influencing the level of experienced conflict.

The focus of our study was to determine how simple effects (congruence, instruction type, believability) interact and effect response times and confidence levels. We predicted participants would experience a greater level of conflict when base rates and stereotypes do not point toward the same answer. This effect should be stronger when instructed to reason based on mathematical probability compared to reasoning based on belief. The level of experienced conflict is expected to increase the likelihood of Type 2 reasoning and manifest as prolonged response times and lowered judgements of confidence. We predicted the addition of different believability levels of stereotypes would influence judgements and response times differently depending on the type of instruction. When reasoning according to belief high believability should mitigate the influence of conflict while increasing it when reasoning according to mathematical probability.

Method

Participants and design

The sample (N=38) was recruited among undergraduate psychology students. The experiment was a 2(instruction type) x 2(congruence) x 2(believability) repeated measures design.

There were two different instructions. For half of the items participants were instructed to answer according to their belief. The exact instruction was “For the following items respond as fast as you can according to what you know to be probable from everyday experience”. For the other half of the items they were instructed to respond in accordance with mathematical probability. The exact instruction was “For the following items respond as fast as you can according to mathematical probability”.

The second independent variable varied the level of congruence. For half of the items there was no conflict between belief and mathematics based probability while for the other half there was.

Finally we varied the level of believability of the content by creating two types of items as displayed in the materials section.

In order to control order effects half of the participants first completed the block based on belief instructions and then the one based on mathematical probability, while the other half had the reversed order. Items within each block were randomized for each participant.

Materials

Previous research with base rate tasks regularly used extreme base rate ratios to achieve the conflict effect, 995/5 and higher (De Neys & Glumicic, 2008; De Neys, Vartanian, & Goel., 2008; Obrecht & Chesney, 2016; Pennycook et al, 2014). We chose to use less extreme and random ratios in order to control for a possible effect of repetitiveness on participant decisions. We implemented clear rules for ratio selection. The highest allowed ratio was 900/100, while the lowest was set at 850/150. Within this spread ratios were generated randomly by a computer algorithm.

To vary believability we followed two different approaches. For half of the items, attributes associated with the person were characteristic of one group, but not exclusive to it. E.g. *Person A is physically attractive. Person A is chosen from a group of supermodels or a group of secretaries.* Obviously, secretaries can be attractive, but the attribute is an integral part of the concept *supermodel*. For the other half, we increased the difference in the believability of the attribute for the two groups. For one group the attribute was still integral, but for the other it was highly uncharacteristic, the opposite of what people would expect. E.g. *Person B is courageous. Person B is chosen from a group of firefighters or a group of deserters*. Courage is an integral part of what people think of firefighting, while cowardice is strongly associated with deserting.

Conflict is achieved when base rate ratios and belief based probability do not point towards the same choice. By adhering to the before mentioned guidelines, forty main and four practice items were selected from a larger pool of constructed items based on researcher scores. Base rate ratios were assigned randomly to the items which were then randomly assigned to experimental conditions. Examples for the four possible combinations of believability and congruence can be seen in Table 1. When combined with the two different instruction types this forms a total of eight experimental conditions (five items per condition).

Table 1: Examples of item types.

	Attribute	Subgroups
Lower believability/Congruent	Elegant	854 ice skaters, 146 teachers
Lower believability/Conflict	Creative	866 waiters, 134 painters
Higher believability/Congruent	Comical	880 comedians, 120 morticians
Higher believability/Conflict	Honest	842 smugglers, 158 judges

The order in which subgroups appeared was randomized among items to avoid habitual responses from our participants.

Procedure

The experiment was designed in E-Prime v2.0.10.356 and conducted in the Laboratory for Experimental Psychology. Before the main experiment participants underwent practice to associate themselves with the way in which they were required to react. For the main experiment participants were told an attribute describing a person would be presented for a few seconds after which they would receive information about the groups from which the person was randomly selected. Finally, they were presented with a choice and were required to answer from which group they thought the person was probably chosen based on one of the instruction criteria (belief in one block, and mathematical probability in the other). Confidence judgements were made on a six point scale, with each point representing a percentage of confidence. Scale value 1 represented 50% confidence (guessing) with each successive value representing an increase of 10% with the scale value 6 representing 100% confidence (complete confidence). An example of the single trial procedure can be seen in Figure 1.

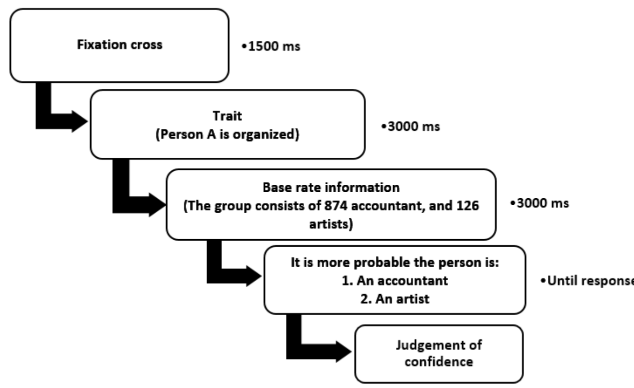

Figure 1: Example of a single trial.

Results

Prior to analysis response time data was processed to eliminate outliers by removing responses outside of the +/- 3 standard deviation range. Outliers made up of 2.2% of all responses. Response times were averaged for items within each experimental condition (5 items per condition) for final analysis. Confidence ratings were also averaged to get the final confidence judgements for each condition. Before the main analysis two 2x2x2x2 mixed analyses of variance were conducted with an additional variable of block order (instruction order) to determine whether the order in which participants completed the experiment influenced response times and confidence judgements. In both analyses the main effect of block order and interactions which include the effect were not significant. The results of these analyses show that the order in which participants completed the task had no influence on response times and confidence judgements.

A 2 (instruction condition) x 2 (congruence) x 2 (believability) repeated measures analysis of variance was conducted on response time data. Results of the analysis can be seen in Table 2. A strong main effect of congruence showed response times were significantly shorter for congruent compared to conflict trials. Higher believability in general led to slower responses, but this was due to a strong effect in conflict situations when participants were instructed to reason according to probability (see the three-way interaction interpretation). The main finding of this study is reflected in the significant two-way interaction between instruction type and congruence which is shown in

Figure 2 (for all figures error bars represent 95% confidence interval for the mean).

Table 2: ANOVA results for response times.

Effect	F (1, 37)	η_p^2
Instruction	.05	.00
Congruence	11.68**	.24
Believability	7.07*	.16
Instruction by Congruence	7.70**	.17
Instruction by Believability	2.32	.06
Congruence by Believability	2.69	.07
Three-way interaction	11.34**	.23

*p<.05; **p<.01

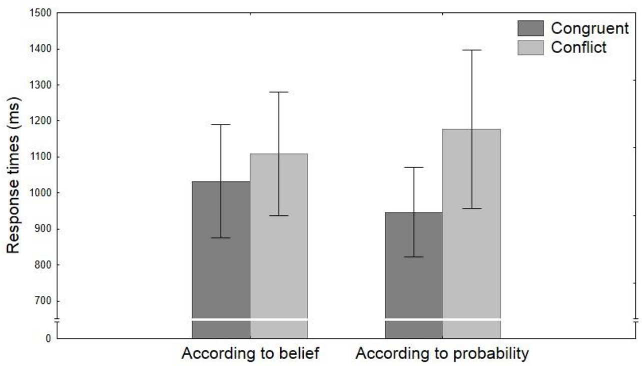

Figure 2: Response times as a function of conflict and instruction type.

The influence of conflict on response times was significantly lower when participants responded in accordance with belief (mean difference between congruent and conflict responses M_{diff} = 76 ms) than when they responded in accordance with mathematical probability (M_{diff} = 230.36 ms). The three-way interaction effect reflects the different influence of believability depending on the instruction. Higher believability mitigated the influence of conflict when participants responded according to belief, and increased it when they responded in accordance with mathematical probability.

The same analysis was conducted for confidence judgements for which results can be seen in Table 3. The analysis showed a similar pattern of results. Participants expressed lower levels of confidence for conflict compared to congruent items, which is reflected in a large main effect of congruence. Believability, in general, slightly increased confidence ratings but was present in more complex interaction effects. Again, the main finding of this study is best observed by considering the significant two-way interaction between instruction type and congruence in Figure 3. Conflict lowered confidence judgements for mathematics based reasoning (mean difference between congruent and conflict responses M_{diff} = 5.24%) more than for reasoning based on belief (M_{diff} = 2.38%).

Table 3: ANOVA results for confidence judgements.

Effect	F (1, 37)	η_p^2
Instruction	.21	.00
Congruence	20.49**	.36
Believability	6.54*	.15
Instruction by Congruence	5.63*	.13
Instruction by Believability	.36	.00
Congruence by Believability	2.30	.06
Three-way interaction	4.77*	.11

*p<.05; **p<.01

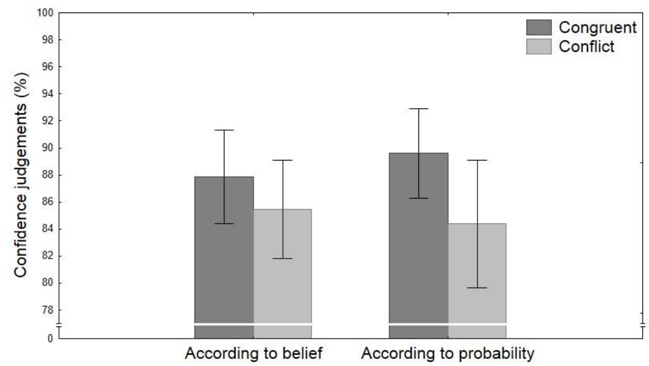

Figure 3: Confidence judgements as a function of conflict and instruction type.

In the three-way interaction believability increased confidence ratings and mitigated conflict influence when participants responded in accordance with belief, and lowered confidence ratings for conflict items when they responded in accordance with mathematical probability.

Next, we analyzed accuracy depending on instruction type. For the belief based instruction the stereotypical response could be considered as correct, and for the probability instruction the opposite. Accuracy was high for the two belief conditions and the congruent probability condition (above 86%), but low for conflict trials in the probability condition (46.84%). Since this data was not distributed normally, we tested differences using the Wilcox matched pairs test. Instruction to reason according to mathematical probability significantly lowered stereotypical responses in conflict trials (Z = 4.94, p<.01), the same was true for the belief instruction condition (Z = 2.65, p<.01). To evaluate confidence judgements we calculated differences between confidence levels and accuracy for each participant (for this and similar procedures see Koriat, Lichtenstein, & Fischoff, 1980). An instruction by congruence ANOVA showed a significant interaction effect (F(1, 34) = 34.81, p<.01, η_p^2 = .51). Results showed the difference between confidence and accuracy was largest for conflict trials in the probability instruction condition (Figure 4).

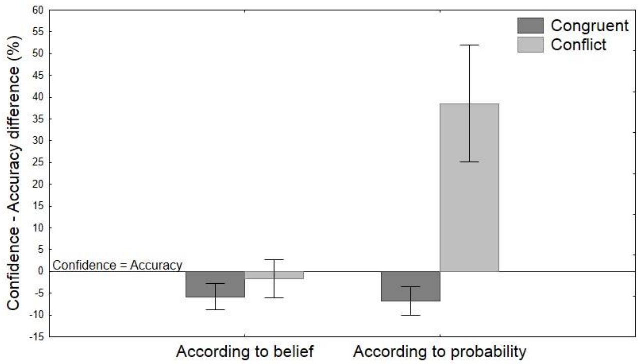

Figure 4: Differences between confidence and accuracy as a function of instruction type and congruence

Additionally, analyses of variance were conducted only for responses that were correct depending on the instruction. A total of 25 participants made up this dataset while the rest did not have correct responses for all of the experimental conditions. For both response times and judgements of confidence, results followed a very similar pattern to the analysis of the full dataset. Once again participants were considerably slower for conflict items ($F(1, 23) = 12.71$, $p<.01$, $\eta_p^2 = .36$), and slightly slower for more believable items ($F(1, 23) = 9.63$, $p<.01$, $\eta_p^2 = .29$). The instruction type by congruence interaction ($F(1, 23) = 9.54$, $p<.01$, $\eta_p^2 = .29$) was again the most interesting result. Conflict had a greater influence on response times when reasoning according to probability than when reasoning according to belief. The three-way interaction was no longer significant but showed the same pattern of results. For judgements of confidence the analysis showed significantly lower levels of confidence in conflict compared to congruent conditions ($F(1, 23) = 19.60$, $p<.01$, $\eta_p^2 = .46$). The instruction by congruence interaction remained significant ($F(1, 23) = 5.62$, $p<.05$, $\eta_p^2 = .20$) and showed conflict lowered confidence judgements to a larger extent than when reasoning according to probability. Additionally, instruction by congruence ($F(1, 23) = 12.37$, $p<.01$, $\eta_p^2 = .35$) and congruence by believability ($F(1, 23) = 4.92$, $p<.05$, $\eta_p^2 = .18$) interactions showed participants were less confident for high believability items when reasoning according to probability and that conflict had a larger influence on confidence for higher believability items. The three-way interaction was no longer significant, but showed the same pattern of results as the analysis of the total response data.

Finally, we calculated an item-level correlation between response times and confidence ratings. Results ($r(38) = -.56$, $p<.01$) showed that participants gave higher judgements of confidence for items they responded to faster.

Discussion and conclusions

According to the proposed dual-process model by Pennycook et al. (2015), initial responses are generated by Type 1 processes in reasoning tasks. If there is a conflict between the initially generated responses, and it is successfully detected, Type 2 processing resolves the conflict in two possible ways. One outcome is the acceptance of a dominant initial response (rationalization), and the other is choosing an alternative response (cognitive decoupling). Because of the expected dominance of belief based responses, we predicted that induced conflict would have a greater influence when instructed to reason based on mathematical probability compared to reasoning based on everyday belief. According to the prediction, this greater influence would initiate Type 2 processes to a larger extent, which would manifest in prolonged response times and lower confidence judgements in the mathematical instruction condition. Both three-way ANOVAs (Tables 2 and 3) prove this prediction to be correct. The expected strong main effect of congruence was significant, which is the usual result in this type of research (Pennycook et al., 2015; Pennycook et al., 2014; Thompson & Johnson, 2014; Thompson et al., 2013). The main findings show that participants responded slower in conflict trials when reasoning according to mathematical probability. Conflict influence was less prominent when reasoning in accordance with everyday belief. This pattern of results is evident for both response times and confidence judgements (Figures 2 and 3). We hypothesize that stereotypical responses have a greater weight during initial response generation (Type 1 processing), which leads to stronger interference of belief on probability based reasoning than vice versa.

Our additional experimental manipulation of stereotype believability resulted in significant three-way interaction effects (Tables 2 and 3). When reasoning according to belief, higher believability mitigates the impact of conflict on response times and confidence levels. On the other hand, higher levels of believability increase the influence of conflict when reasoning based on mathematical probability. This result may represent further proof for the existence of differently weighted initial responses among which belief based responses are very prominent.

Participants expressed a higher level of confidence for items which had shorter response times indicating response fluency is a strong cue in the formation of metacognitive judgements. This finding was obtained in recent studies using different thinking and reasoning tasks (see Thompson et al., 2013).

Furthermore, when instructed to reason according to belief, conflict decreased stereotypical responses, but to a far lesser degree than when instructed to reason according to mathematical probability. It is important to note that even when instructed to reason according to probability participants chose belief based responses in over 50% of conflict trials. This further strengthens the conclusion that everyday belief dominates reasoning in this specific task. Within the framework proposed by Pennnycook et al. (2015), this would indicate instruction to reason according to probability influenced the relative importance of belief and probability information, but did not fully override the initial dominance of belief based responses. Participants were overconfident only for conflict trials in the probability instruction condition (Figure 4). This was probably due to the fact that in the other three conditions the dominant belief based answers were correct, while in this one that was not the case.

When we analyzed only correct responses the same pattern emerged. As the model predicts, participants were slower and less confident in conflict trials and the conflict had a larger effect when reasoning according to probability. Based on these results, we can conclude that emphasizing a particular way of reasoning can have an effect on the relationship between conflict, response fluency and metareasoning judgements.

When these results are considered together we can conclude everyday belief has a stronger interference on mathematics based reasoning in this type of task than vice versa. Since a main effect of instruction (reasoning type) was not observed, we can speculate the two processes run in parallel, but that the result of the belief based process has a higher weight.

The results may have practical implications, particularly in educational settings. Many tasks require students to ignore intuitive modes of reasoning in favor of analytical thinking, and it is in those types of tasks where results such as found by this study could be applied to increase efficiency.

To conclude, the results of this study confirm the strong influence of conflict on response times and confidence levels in reasoning tasks. The study expands on previous research by introducing further complexity into established relationships between processes. Explicit instructions in combination with different levels of believability moderate the influence of conflict on fluency and confidence judgements. Results may indicate parallel processing of multiple, differently weighted processes, but more sophisticated research is required to explore the findings further.

Acknowledgments

This research was supported by Grant 4139 from the Croatian Science Foundation, and the University of Rijeka Research grant 13.04.1.3.11.

References

Ackerman, R., & Thompson, V.A. (2015). Meta-reasoning: What can we learn from meta-memory? In A. Feeney & V.A. Thompson (Eds.), *Reasoning as Memory* (pp. 164-182), New York: Psychology Press.

Evans, J.St.B.T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning, 13*(4), 321-339.

Evans, J.St.B.T. (2012). Questions and challenges for the new psychology of reasoning. *Thinking & Reasoning, 18*(1), 5-31.

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

Evans, J.St.B.T., Handley, S.J., Neilsen, H., & Over, D. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *The Quarterly Journal of Experimental Psychology, 63*(5), 892-909.

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning, 20*(2), 169-187.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*, 1248-1299.

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.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237-251.

Koriat, A., Lichtenstein, S., & Fischoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 107-118.

Markovits, H., Thompson, V.A., & Brisson, J. (2015). Metacognition and abstract reasoning. *Memory & Cognition, 43*, 681-693.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Brower. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 26, pp. 125-173). San Diego: Academic Press.

Obrecht, N.A., & Chesney, D.L. (2016). Prompting deliberation increases base-rate use. *Judgment and Decision Making, 11*(1), 1-6.

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.

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.

Thompson, V.A., Evans, J.St.B.T., & Campbell, J.I.D. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning, 19*(3), 431-452.

Thompson, V.A., & Johnson, S.C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning, 20*(2), 215-244.

Thompson, V.A., Turner, J.A.P., Pennycook, G., Ball, L.J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition, 128*, 237-251.

Trippas, D., Thompson, V.A., & Handley, S.J. (2016). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory & Cognition*. Advance online publication. DOI: 10.3758/s13421-016-0680-1.