When Intuitions Are Helpful: Prior Beliefs Can Support Reasoning In The Bat-And-Ball Problem

Bastien Trémolière
Cognition, Langues, Langage, Ergonomie
Université de Toulouse

Wim De Neys
LaPsyDE (Unité CNRS 3521)
Université Paris-Descartes

Intuitions are often considered suboptimal, because they can bias people's thinking. The bat-and-ball problem is a celebrated example of this potentially detrimental aspect of intuitions since it elicits a very appealing and prepotent intuitive but incorrect response. We propose to show that certain kinds of intuitions (i.e., prior beliefs) can help people to reason better on this task. In two experiments, participants answered either a classic congruent version of the bat-and-ball problem in which the intuitively cued response fitted with prior knowledge (i.e., was believable) or a modified incongruent version in which the intuitively cued response conflicted with prior knowledge (i.e., was unbelievable). Results indicate that participants who solved the modified unbelievable version performed better than participants who solved the classic believable version. Our data highlight that prior beliefs, even in the bat-and-ball problem, can accidentally make people perform better, probably because they encourage them to adopt a more effortful processing strategy.

Keywords:. Bat-and-Ball Problem; Prior Beliefs; Effortful Reasoning

Intuitions have long been considered as suboptimal, because they can lead to incorrect outcomes and bias people's thinking (Evans, 2008; Kahneman, 2011). In this perspective, many ingenious tasks have been designed that cue a salient intuitive, but incorrect answer. One of the most striking and famous examples is the so-called bat-and-ball problem (Bourgeois-Gironde & Van Der Henst, 2009; Frederick, 2005). The classic version of this problem is usually framed as follow:

> A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

Interestingly, the response that intuitively springs to people's mind is '10 cents'. However, although this answer is particularly appealing, it is wrong. Saying that the ball costs $.10 would require that the bat costs $1.10. Therefore, the total cost would no longer be $1.10, but now $1.20; hence, the bat cannot cost $.10. The correct answer is that the ball costs $.05 (and the bat costs $1.05). A possible explanation which can account for this error is that people seem to substitute the 'more than' relational statement by an easier absolute statement: the sentence is no longer read as 'The bat costs $1.00 *more than* the ball' but now it is read as 'The bat costs $1.00'. That is, people parse $1.10, into $1 and $.10, which is easier than working out the sum. Specifically, this intuitive response would seem to result from a process that is known as attribute substitution: when people face a difficult question, they tend to replace it with an easier one (see Kahneman & Frederick, 2002; Kahneman, 2011).

Critically, research that shows a supportive role of intuitions in reasoning is more scarce. Although few authors would deny that intuitions can sometimes be valuable, the field of the psychology of reasoning has tended to capitalize on situations or tasks where intuitive considerations conflict with reflective ones (such as in the bat-and-ball problem). Hence, a possible supportive role of intuitions in reasoning has received less attention (e.g. Gladwell, 2005; Gigerenzer, 2007; Gigerenzer, Todd, & the ABC Research Group, 1999).

In this paper we present evidence for the supportive role of intuitions in the bat-and-ball problem. We thereby focus on a manipulation of the believability of the intuitively cued '10 cents' response. Note that in the classic problem one could assume that this response is rather neutral, in that it is neither strongly believable nor strongly unbelievable[1]. Contrast now this with the following modified version:

> A Ferrari and a Ford together cost $190000. The Ferrari costs $100000 more than the Ford. How much does the Ford cost?

This modified incongruent problem has the exact same structure as the original version. People will read the 'Ferrari

[1] Clearly, one might argue about how reasonable a 10 cents price for a ball really is. It might be somewhat cheap when buying a new ball but it might be fairly appropriate as a price that is paid when practicing at a batting cage, for example. Our point is that our study will directly manipulate the believability of the cued intuitive response. Note that the believability of the cued intuitive response in the problem versions that we adopted was also explicitly pretested (see further).

costs 100 000 more' as 'the Ferrari costs $100 000' which will cue the intuitive answer '$90 000'. However, this intuitively cued response will be far less believable. That is, our background knowledge tells us that a Ford is but an ordinary car. Hence, given this knowledge people should find it a bit bizarre and somewhat unlikely that a Ford costs '$90000'. We hypothesize that such a conflict with our background knowledge will decrease the appeal of the substituted response and might thereby actually help people to reason better. We tested this basic idea in two experiments.

Pretest

Our reasoning is that the manipulation of the believability of the intuitively cued response will shape participants's performance, with greater performance for participants who face unbelievable prices (in our case, 'a Ford costs $90000).

To ensure that the price reflecting the intuitively cued response ($90000) is consistent with people's prior beliefs, we conducted a pretest in which 47 independent american participants (36 women, mean age = 41.4 years, SD = 12.8) indicated on a 10-point scale (from 0, not at all believable, to 10, totally believable) whether $90000 sounded like a believable price for a Ferrari, for a Ford and for a Rolls Royce. Clearly, results indicated that $90000 is a believable price for a Ferrari (mean = 7.1, SD = 2.7) and for a Rolls Royce (mean = 6.8, SD = 2.8), but is widely unbelievable for a Ford (mean = 2.5, SD = 2.9). Paired samples t-tests detected significant differences between the believability of the price of the Ferrari and that of the Ford, $t(46) = 9.4$, $p < .001$, as well as between the believability of the price of the Rolls Royce and that of the Ford, $t(46) = 7.7$, $p < .001$. No difference was detected between the believability of the price of the Ferrari and that of the Rolls Royce, $t(46) = 0.7$, $p = .49$.

Because men are usually more knowledgeable about car prices than women and because prior knowledge mediates judgment about automobile price (Herr, 1989), we conducted a second stage of analyses in which gender was entered as predictor. The analyses detected a significant difference between men (mean = 4.2, SD = 3.3) and women (mean = 2.0, SD = 2.5) when evaluating the believability of the price of the Ford, $t(45) = 2.4$, $p = .023$. No gender difference was detected for the Ferrari and for the Rolls Royce.

Experiment 1

Method

167 american participants (99 women, mean age = 31.9 years, SD = 12.1) were recruited on the MTurk online platform. Each participant completed one of the two versions of our bat-and-ball problem (classic version or modified version).

Participants in the classic condition (in which the cued response response is believable) saw the following problem:

> A Ford and a Ferrari together cost $190000. The Ford costs $100000 more than the Ferrari. How much does the Ferrari cost?

In this example, prior beliefs (that is that a Ferrari is expensive) would fit with the cued response in that they both lead to conclude that a Ferrari costs $90000 and not a cheaper price. Participants in the incongruent version (in which the cued response is unbelievable) saw the problem presented in the introduction ('How much does the Ford cost?').

Results

As expected, participants performed better on the modified incongruent version than on the classic congruent version. 27% ($SD = 4\%$) of participants correctly solved the modified incongruent version while only 11% ($SD = 3\%$) of participants managed to solve the classic version. These descriptive statistics were supported by a χ^2 analysis which detected a significant association between the condition and the performance, $\chi^2(1, N = 167) = 6.58$, $p = .01$.

Note that we also looked at potential differences in the kind of errors made as a function of the condition. That is, we contrasted whether participants who erred gave the intuitively cued response or 'other' incorrect responses (i.e., all incorrect responses that are not the intuitively cued incorrect response). The analysis detected no difference in the kind of errors made, with the frequency of 'other' errors being very low (congruent version, 16.7% of the errors; incongruent version, 11.1% of the errors), $\chi^2(1, N = 135) = 0.86$, $p = .35$.

As the believability of the price of the Ford slightly differed for males and females in the pretest, we looked at potential gender differences in task performance for the two versions of the task. The χ^2 analyses detected no difference between males and females either for the congruent version ($p = .10$) or the incongruent version ($p = .28$).

Results confirm our hypothesis that prior beliefs can help people to reason better to the bat-and-ball problem, when they conflict with the cued response. Nevertheless, a limitation in the design of our problems is that the classic congruent version of the task also elicits a conflict between prior beliefs and the truth of the statement presented in the instructions (which has to be considered as true for the purpose of the experiment). Specifically, it is probably counterintuitive for a majority of people to accept that a Ford is more expensive than a Ferrari, as it is stated in the problem ('The Ford costs $100000 more than the Ferrari'). In a second experiment, we aim to consolidate the results we obtained in Experiment 1, while designing a more rigorous congruent version of the task.

Experiment 2

Method

222 american participants (124 women, mean age = 32.6 years, SD = 11.5) were recruited on the MTurk online platform. As in Experiment 1, each participant completed one of the two versions of our bat-and-ball's type problem (classic congruent version or modified incongruent version).

The modified incongruent version of the task is the same as that used in Experiment 1 (The Ferrari costs $100000 more than the Ford). We now display the new congruent version of the task:

> A Rolls-Royce and a Ferrari together cost $190000. The Rolls-Royce costs $100000 more than the Ferrari. How much does the Ferrari cost?

In this example, we fixed the problem mentioned above, by eliminating the conflict between prior beliefs (about the plausibility of the prices of the Rolls-Royce and the Ferrari) and the statement in the instructions ('The Rolls-Royce costs $100000 more than the Ferrari').

Results

As in Experiment 1, participants performed better on the modified incongruent version than on the classic congruent version. 25% ($SD = 4\%$) of participants correctly solved the incongruent version while only 14% ($SD = 3\%$) of participants managed to solve the congruent version. These descriptive statistics were supported by a χ^2 analysis which confirmed this difference, $\chi^2(1, N = 222) = 4.0, p = .054$.

As in Experiment 1, we looked at potential differences in the kind of errors made as a function of the condition. The analysis detected no difference between the two versions, $\chi^2(1, N = 135) = 0.86, p = .35$, with the frequency of 'other' incorrect responses being critically low (congruent version, 10.3% of the errors; incongruent version, 11.1% of the errors).

We also looked at potential gender differences for the two versions. The χ^2 analyses detected no difference either for the congruent version ($p = .10$) or the incongruent version ($p = .78$).

Finally, Experiment 2 replicates the results of Experiment 1, by showing that participants who solved the modified incongruent version of the task performed better than participants who solved the congruent version of the task.

Discussion

In this article, we aimed to show that some aspects of intuitions could support reasoning and could help people to reason better in certain contexts. We ran two experiments in which participants performed either a congruent, believable version (i.e., the intuitively cued response fitted with background knowledge) or an incongruent, unbelievable version (i.e., the intuitively cued response conflicted with background knowledge). We predicted and experimentally observed that people facing the modified incongruent version were more likely to solve the problem correctly than participants who faced the congruent version.

What these data tell us is that in some cases, such as when different intuitions conflict with each other (as it is the case here, since in the modified incongruent version background beliefs conflict with the traditional cued intuitive response), some of these intuitions can support deliberative reasoning, probably by decreasing the salience of other kinds of intuitions. Arguably, this could encourage people to engage in more reflective thinking. But how would it work?

A first explanation would directly refer to the selective processing account of reasoning, which postulates that people typically follow their intuition when the response is congruent with their beliefs, only to switch to a more effortful reasoning when the response is incongruent with their beliefs (Stupple, Ball, Evans, & Kamal-Smith, 2011)[2]. Interestingly, this account enables one to rule out the issue raised by the (un)believability of the preamble (e.g., the Ford costs $100000 more than the Ferrari), by emphasizing that participants only judge the believability of the intuitively cued response (the Ferrari costs $90000).

An alternative explanation would could come from the fluency framework (for a review, see Oppenheimer, 2008). Briefly, fluency is defined as people's subjective experience of ease of processing of mental actions. Fluency is used as a mechanism for cue selection during decision making and encourages people to switch between automatic processes and more reflective processes. Specifically, fluency would lead people to use an intuitive processing strategy, while disfluency would lead them to use a more effortful processing strategy. Several studies showed that fluency directly influenced a wide range of domains, including reasoning strategy selection (e.g., see Alter & Oppenheimer, 2006; Alter, Oppenheimer, Epley, & Eyre, 2007; Thompson & Morsanyi, 2012; Thompson et al., 2013). In our studies, it seems reasonbale to assume that fluency decreases when prior beliefs conflict with the intuitive response (i.e., modified version). A consequence would be a greater encouragement to engage in more effortful thinking.

Therefore, disfluency in the incongruent, unbelievable version is probably making people perform better by helping them at detecting the conflict between intuition and logic or by helping them, if the conflict is already detected, at inhibiting the traditional intuitively cued incorrect response, by

[2] Although this model is basically applied on syllogistic reasoning problems featuring belief bias, it affords interesting clues to the understanding of the cognitive mechanisms underlying our results.

making it less appealing (see De Neys, 2012; De Neys & Bonnefon, 2013; De Neys, Rossi, & Houdé, 2013).

We readily acknowledge that although our results showed that prior beliefs could improve reasoning, our effects are small: we only observed a 10% accuracy boost. However, given the well-known failures at optimizing reasoners' performance on this problem (e.g., Bourgeois-Gironde & Van Der Henst, 2009) even such a small increase is far from trivial. That being said, it is clear that the key contribution of the present paper lies in the demonstration of the basic principle. Obviously, future work should focus on how this factor can be fine-tuned to increase its impact, in particular in situations in which people are less likely to adopt an effortful processing strategy.

The bat-and-ball problem is known to be very difficult, because it strongly leads people to rely on intuitions to give an easy, automatic but incorrect answer. Here we showed that making different types of intuitions (i.e., intuitions cued by prior beliefs and the intuitively cued substituted response) conflict can be beneficial for reasoning. We believe that this helps to highlight the sometimes neglected supportive role of intuitions for human thinking.

References

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences*, *103*(24), 9369-9372.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: metacognitive difficulty activates analytic reasoning. *Journal of experimental psychology. General*, *136*(4), 569-576.

Bourgeois-Gironde, S., & Van Der Henst, J.-B. (2009). How to open the door to system 2: Debiasing the bat-and-ball problem. In S. Watanabe, A. P. Bloisdell, L. Huber, & A. Young (Eds.), *Rational Animals, Irrational Humans* (p. 235-252). Tokyo: Keio University Press.

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*, 28-38.

De Neys, W., & Bonnefon, J.-F. (2013). The 'whys' and 'whens' of individual differences in thinking biases. *Trends in cognitive sciences*, *17*, 172-178.

De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, *20*, 1-5.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255-278.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25-42.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. London: Pengiun.

Gigerenzer, G., Todd, P., & the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Gladwell, M. (2005). *Blink: The power of thinking without thinking*. New York: Little, Brown.

Herr, P. M. (1989). Priming price: Prior knowledge and context effects. *Journal of consumer research*, *1*, 67-75.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus & Giroux.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (p. 49-82). Cambridge: Cambridge University Press.

Stupple, E. J. N., Ball, L. J., Evans, J. S. B. T., & Kamal-Smith, E. (2011). When logic and belief collide: Individual differences in reasoning times support a selective processing model. *Journal of Cognitive Psychology*, *23*(8), 931-941.

Thompson, V. A., & Morsanyi, K. (2012). Analytic thinking: do you feel like it? *Mind and Society: Cognitive Studies in Economics and Social Sciences*, *11*(1), 93-105.

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