Individual differences in belief bias: Inhibition leads to better reasoning but also to more confirmation bias

Gabriela Iwama; Víthor Franco; Luciano Buratto

University of Brasília, Brazil

gabriela.yukari.iwama@gmail.com

Introduction

Syllogism	Conclusion: Believable	Conclusion: Unbelievable
Valid	No cigarettes are inexpensive. Some addictive things are inexpensive. Therefore, some addictive things are not cigarettes. P("Valid") = .92	No addictive things are inexpensive. Some cigarettes are inexpensive. Therefore, some cigarettes are not addictive. P("Valid") = .46
Invalid	No addictive things are inexpensive. Some cigarettes are inexpensive. Therefore, some addictive things are not cigarettes. P("Valid") = .92	No cigarettes are inexpensive. Some addictive things are inexpensive. Therefore, some cigarettes are not addictive. P("Valid") = .08

Table 1. *Endorsement rates for believable and unbelievable syllogism according to their logical validity. Adapted from Klauer, Musch and Naumer (2000).*

- In solving syllogisms, individuals tend to endorse believable conclusions regardless of their logical validity. Endorsement of unbelievable conclusions, however, depends on logical validity (belief bias effect).
- Usually interpreted as better reasoning for unbelievable conclusions, recent studies using Signal Detection Theory support the view that belief bias reflect a change in response bias[1].
- Individuals with higher levels of inhibitory control (measured by Cognitive Reflection Test) show less belief bias[2].

Research Goals

1. **Test individual differences in inhibitory control (Cognitive Reflection Test), while controlling for Numeracy skills**
2. **Test the influence of logical validity on belief judgments**

Method

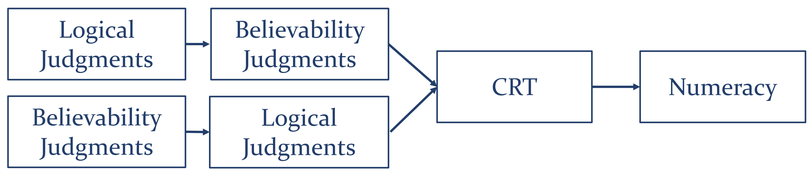
Participants.
124 volunteers participated of an online study conducted on Psytoolkit[3].

Measures

Syllogisms with Belief Bias[1]. Discriminability between valid and invalid syllogisms, $d' = \Phi^{-1}(Hit) - \Phi^{-1}(FA)$, and the response bias, $c = -\frac{\Phi^{-1}(Hit)+\Phi^{-1}(FA)}{2}$, were calculated for 64 syllogisms with Believable and Unbelievable conclusions.

Cognitive Reflection Test (CRT)[4]. 6 questions about simple math problems that induce wrong intuitive responses.

Berlin Numeracy Test[5]. 4 multiple choice probability problems.



Path Analysis

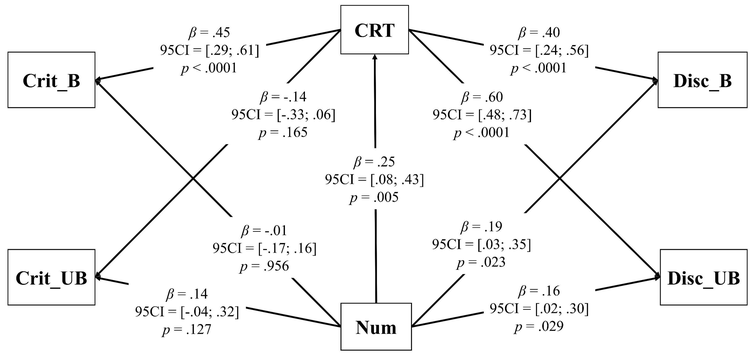


Figure 1. Path analysis showed that CRT scores mediated the effect of Numeracy (Num) on the discriminability of believable (Disc_B) and unbelievable (Disc_UB) syllogisms, as well as on the criteria of believable (Crit_B) and unbelievable (Crit_UB) syllogisms.

Order effects

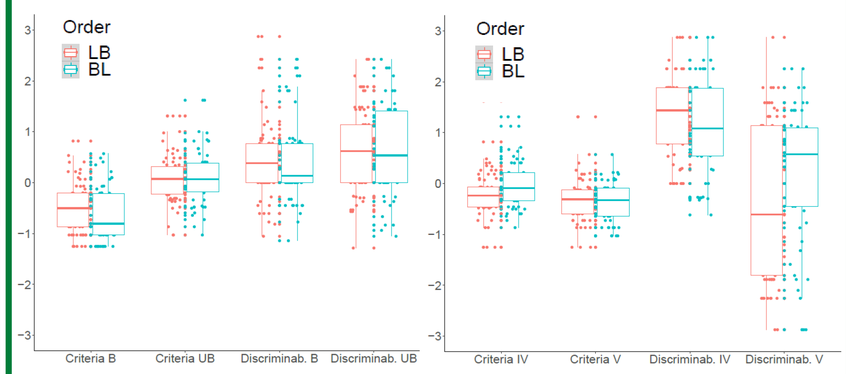


Figure 2. Discriminability and criteria for logical (left) and belief judgments (right) according to presentation order (LB: logical judgments followed by believability judgments; BL: believability followed by logical judgments).

Cognitive Reflection Test

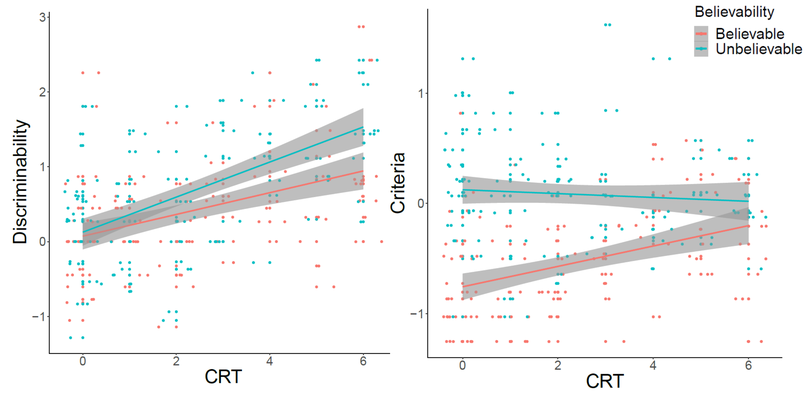


Figure 3. Discriminability (left) and criteria (right) for syllogisms with believable and unbelievable conclusions as a function of CRT scores.

Numeracy

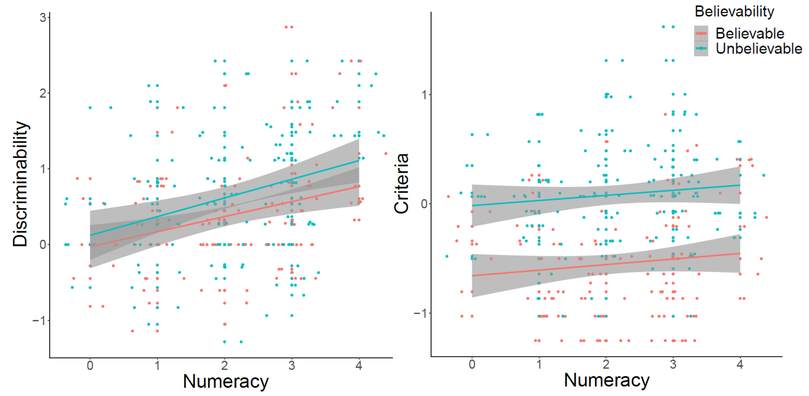


Figure 4. Discriminability (left) and criteria (right) for syllogisms with believable and unbelievable conclusions as a function of Numeracy.

Conclusion and Future Directions

1. Individuals with higher scores on the CRT, a measure of inhibitory control, were better able to solve syllogisms with unbelievable conclusions than individuals with low scores on the CRT. Those individuals also showed more confirmation bias (endorsement of believable conclusions).
2. We replicated and extended Trippas et al.'s (2018) findings by showing that performance on the CRT predicted discriminability of believable and unbelievable syllogisms even after controlling for numerical skills (Numeracy score).
3. The finding that CRT scores had a greater impact on unbeliavable syllogisms might be explained by motivated cognitive control[6] (e.g., higher probability of engaging in deliberative processes when disconfirming conclusions).

References

1. Dubé, Rotello & Heit (2010) 2. Trippas et al. (2018) 3. Stoet (2017) 4. Primi et al. (2016) 5. Cokely et al. (2012) 6. Botvinick & Braver (2015)