Analyzing the Moses Illusion

A Data-Driven Study of Semantic Errors

Giorgia Sorrentino

Introduction

- The Moses illusion is defined as a phenomenon in which individuals fail to detect inconsistencies in a question [1].
- A famous example is “How many pairs of each animal did Moses bring on his ark?”. Many participants to the 1981 study answered the question by simply saying “two”.
- This question represents an illusion, since the biblical story mentioned that Noah was the individual who built the ark and brought the animals aboard, not Moses.
- Currently, there is not a unique explanation about the existence of this phenomenon.
- A possible hypothesis explored so far is the truth bias, which can be described as individuals not expecting to be intentionally misled in a communication [2]).

Methods

A total of 54 participants were included in the experiment. They were presented with a series of general knowledge questions.

For the purpose of reproducibility and privacy, this poster presents analyses on a synthetic dataset generated to mimic the structure and statistics of the original Moses illusion experiment. No real participant data was used.

Research Goal

→ The aim was to test the participants’ memory and reading comprehension, as to examine whether they would detect semantic distortions in the questions or accept them as plausible.

Research Question

→ Did the participants fall for the illusion, or did they know the answer?

Materials

The following table shows the 4 different types of question included in the experiment, which were referred to as groups or conditions.

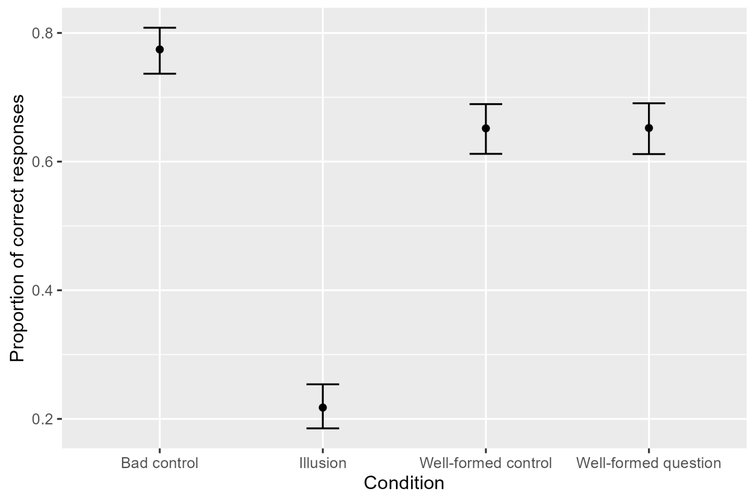
Condition	Example
Illusion question	According to the Bible, how many animals of each kind did Moses take on the ark?
Well-formed question	According to the Bible, how many animals of each kind did Noah take on the ark?
Well-formed control	What is the name of the current chancellor of Germany?
Bad control	Which Nordic country are coconut trees native to?

Procedure

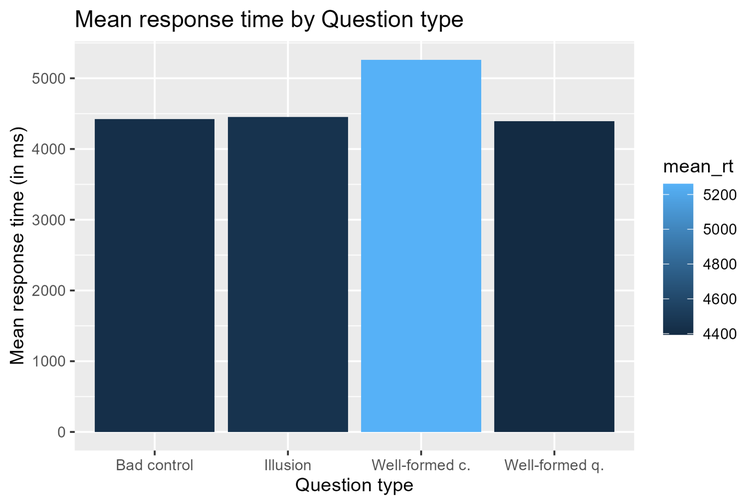
- Each participant was asked a total of 41 questions.
- The collected data was analyzed using the programming language R in the software RStudio [4].
- The tidyverse package was used to analyze the data [5].
- In this synthetic dataset, preprocessing was minimal since the data is already clean. In the real data, steps like duplicate removal, outlier detection and handling missing values would be necessary.

Analysis and results

- All analyses presented ehre use synthetic data generated to reflect typical patterns of the Moses illusion experiment.
- Multiple analyses were conducted to address the research question posed by the experiment.
- The first analysis examined the proportion of correct responses by question type.
- Illusion questions showed the lowest accuracy ($\approx 22\%$, 95% CI 18–25%), while well-formed and control questions were higher (≈ 65–75%).
- Error bars in the plot represent 95% confidence intervals.



- A Chi-square test confirmed that accuracy differed significantly across conditions.
- Logistic regression indicated that the probability of a correct response varied by condition, with illusion questions strongly reducing the likelihood of a correct answer relative to the reference category.
- Another analysis conducted was centered on the response times of the participants, as shown in the graph below.



- Mean response times differed across question types. Well-formed questions were answered fastest (average 4394 ms), while well-formed control questions (average 5261 ms) were slowest.

Conclusions

- The results indicate that the participants were affected by the Moses illusion.
- This illusion can be seen as a clear demonstration of how the human mind prioritizes efficiency and adaptability [3].
- Variation in the mean response times across question types can reflect the different cognitive processing demands.
- Illusion questions showed intermediate response times, indicating that participants often produced answers without extensive reflection, consistent with shallow processing.
- Together with the low accuracy for illusion questions, this pattern suggests that participants frequently failed to detect the anomaly and relied on intuitive but incorrect responses.

Note: This analysis uses synthetic data for illustration purposes.

⇒ The findings of this study align with previous research, suggesting that individuals often process familiar-looking questions in a more superficial or shallow way.

References

[1] Thomas D. Erickson and Mark E. Mattson. From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, 20(5):540–551, October 1981.

[2] Timothy R. Levine. Truth-default theory (tdt): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33(4):378–392, May 2014.

[3] Heekyeong Park and Lynne M. Reder. page 287–304. Psychology Press, December 2012.

[4] Posit team. *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA, 2025.

[5] Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, and McGowan. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686, 2019.