Eberhard Karls Universität Tübingen

Seminar für Sprachwissenschaft

Metaphors and Reliability in Reasoning

Research Apprenticeship

Winter term 2022/2023

Miriam Schiele

Matriculation number: 6098290

Supervised by

Prof. Dr. Michael Franke

Date of submission: March 31st, 2023

Table of contents

Table of figures

Abstract

To be inserted here

1. Introduction

Metaphors can be found in everyday speech. Their persuasive nature has been object of studies since the last 1970s (cf. Gibbs 2010). For instance, Thibodeau & Boroditsky (2011) researched how even subtle metaphors impact decision-making regarding a socio-political issue, namely crime. In their study, they explored how metaphorically framing crime as a *beast* or a *virus* influences the suggestions offered by the participants to solve the issue. That is, when crime was metaphorically described as a *beast*, participants were more likely to suggest enforcement-oriented solutions to the issue than when crime was metaphorically framed as a *virus*.

So far, however, little research has been conducted regarding the question whether the reliability of the speaker influences the effect of the metaphor. Hence, the present study replicates the study conducted by Thibodeau & Boroditsky (2011). The first hypothesis, therefore, is that metaphorically framing crime as a *beast* triggers participants to prefer enforcement-oriented solutions and describing crime as a *virus* influences participants to favor reform-oriented solutions.

In addition, a follow-up study aims to test whether the reliability of the speaker describing the issue influences the effect of the metaphor. Unlike in the replication study, the participants are not only presented with a metaphorical description of the crime issue but also with pictures of different speakers that are associated with different levels of reliability, namely a newscaster and an alcoholic. The second hypothesis, therefore, holds that the effects of both the *beast* and the *virus* frame are higher if the speaker describing the crime issue seems reliable.

The paper is subdivided as follows. Chapter 2 outlines the research in the field of metaphors in reasoning. Chapter 3 focuses on the study design, including the pilot study, the replication study, and the follow-up study. In chapter 4, I will present the results and in chapter 5, I will discuss the findings. Finally, chapter 6 provides a summary of the study as well as an answer to the research question.

2. Literary review

In the following section, I will outline the research in the field of metaphors in reasoning. First, the scientific background of metaphors is explained, including the conditions under which metaphors are most influential. Second, I will summarize the study by Thibodeau & Boroditsky (2011) on metaphors in reasoning and discuss their results. Next, follow-up studies that could replicate the results by Thibodeau & Boroditsky (2011) as well as their further findings are recapitulated. Finally, I will present studies that could not replicate the findings and discuss their results.

Generally, metaphors can be described as "instances of non-literal language that involve some kind of [implicit] comparison or identification" (Knowles & Moon, 2005, p. 7). According to the conceptual metaphor theory, people speak and think in metaphors (cf. Thibodeau & Flusberg, in press). Consequently, recent studies have focused under which conditions and the degree to which metaphors can influence one's thoughts and decisions about socio-political topics such as crime (cf. Thibodeau, Hendricks & Boroditsky, 2017, Thibodeau, Matlock & Flusberg, 2019, Steen, Reijnierse & Burgers 2014, Thibodeau & Flusberg, in press, Reijnierse et al., 2015).

For metaphors to have their greatest impact, several factors must be aligned. First, the source domain (e. g. *virus* or *beast*) and target domains (e. g. *crime*) must match for the metaphor to be apt. Second, people must not already have strong beliefs about the topic in question. Third, prior knowledge of the topic as well as knowledge and interest in the source domain is useful for the metaphor to be persuading. Next, the entailment of metaphors also depends on factors other than the source domain. That is, the meaning of common metaphors might vary. Finally, metaphors are more influential, if the metaphorical mapping is extended in ways that align with the metaphor. For instance, if the language used to describe the metaphor align with the mapping, people more likely to favor the metaphor-consistent response (cf. Thibodeau, Hendricks & Boroditsky, 2016).

However, little research has been conducted so far regarding the question on the influence of the speaker of the metaphor. That is, whether the reliability of the speaker corelates with the extent to which the metaphor is

influential. For instance, if the speaker seems less reliable, the hearer is less susceptible to the metaphor.

In the original study that is to be replicated in the present paper, Thibodeau & Boroditsky (2011) found that metaphors can heavily influence reasoning. In their first experiment, 485 participants were given a short metaphorically framed report about increasing crime rates in the fictitious City of Addison, which included crime statistics of said city. For half of the participants, crime was described as *virus* infecting the city, while for the other half, as a *beast* attacking the city. Afterwards, participants were asked to propose solutions for the crime problem. Participants that were confronted with the metaphor of crime as a *virus* were more likely to propose dealing with the underlying problem and suggesting reforms to prevent crimes. Contrary, metaphorically framing crime as a *beast* led participants to focus on police force or other methods of law enforcement. Interestingly, participants identified the crime statistics in the report as the motivation for their problem-solving decision instead of the metaphor, although the crime statistics were identical in both vignettes. Thus, metaphors seem to subconsciously influence the participants' decisions. Additionally, their results suggest that overall, Republicans are more likely to propose enforcement solutions.

In a replication study, Thibodeau & Boroditsky (2013) asked participants to evaluate a set of problem-solving solutions for the crime problem and select the most effective ones. The results were consistent with those of the original study. Similarly, Thibodeau & Boroditsky (2015), Thibodeau, Iyiewaure & Boroditsky (2015), and Thibodeau (2016) could replicate the findings. However, Thibodeau & Boroditsky (2015) found indications of a cultural shift in the opinion on the topic of crime. That is, participants' suggestions in their study in 2015 were focused on community outreach, while in the original study in 2011, they proposed "policies grounded in the economy and prison system" (p. 11). Additionally, neighborhood watches were no longer seen as reform-oriented and consistent with the *beast* frame, as it was the case in 2011, but were seen as enforce-oriented instead (cf. Thibodeau & Boroditsky, 2015). Moreover, their results suggest that metaphors are more influential if the metaphoric framing is extended into the

texts of policies. That is, participants were more likely to suggest social reforms that were labelled as *treatments* in response to a crime *virus*, while they favored *attacking* the problem with tough enforcement in response to a crime *beast* (cf. Thibodeau, 2016). As in the original study, results of both replication studies showed Republicans, overall, were more likely to be enforcement-oriented (cf. Thibodeau & Boroditsky, 2015, Thibodeau, 2016).

Likewise, Christmann & Göhring (2016) could replicate the findings as well. However, the participants were Germans instead of US citizens, as it was the case in the original study by Thibodeau & Boroditsky (2011). This might have led to different results, since "the use and evaluation of specific metaphors can vary across cultures" (Steen, Reijnierse & Burgers, 2014, p. 7). Additionally, the coding of the answers differed, insofar as that answers that included an equal number of suggestions for enforcement and reform were not counted to either enforcement or reform, respectively, by 0.5, as it was the case in Thibodeau & Boroditsky (2011) but were coded as 0 instead (cf. Christmann & Göhring 2016). Interestingly, when conducting the coding in the same way as in the original study, the difference between the responses of the *beast* and the *virus* group was not statistically significant (cf. ibid.).

The following two studies by Steen, Reijnierse & Burgers (2014) and Reijnierse et al. (2015) could not replicate the findings of the original study either. The results by Steen, Reijnierse & Burgers (2014) failed to show statistically significant differences between the metaphorically frames but suggested that enforcement solutions are preferred overall. However, as pointed out by Thibodeau & Boroditsky (2015), the coding of the answer differed from the original study in a way that is psychologically and statistically problematic. Moreover, only their fourth experiment included a sufficient sample. According to Thibodeau & Boroditsky (ibid.), it would be more appropriate to treat the data as categorial. In their categorical data analysis, they found that participants of the *virus* group favored reform-oriented options that are consistent with the *virus* frame (cf. Thibodeau & Boroditsky, 2015).

Furthermore, in their series of replication studies, Reijnierse et al. (2015) could not find similar evidence as in the original study by Thibodeau

& Boroditsky (2011). That is, they found no significant effect for the metaphorical framing on the preferred measures (cf. Reijnierse et al., 2015, p. 257). However, instead of suggestion solutions, participants were asked to evaluate the effectiveness of two sets of measures on a scale and the presented text of the crime problem severely differed from the one in the original study. Not only did they include the mayor of the city as speaker but also information about the longevity of the crime issue, the length of the report, and the content of the report itself differed. That is, the report did not include statistical information and read *Crime is a virus/beast* instead of *Crime is a virus infecting the city/wild beast preying on the city*, as it was the case in the original study. The authors admitted that the design of the study was inappropriate in such a way that it "may have made it impossible to find out whether people actually reason by working out the entailments of the metaphorical frame" (Reijnierse et al., 2015, p. 260). Yet, as in the original study, the effect that Republicans, overall, were significantly more likely to find the enforcement-oriented approaches more effective than Democrats and Independents could be replicated (cf. Reijnierse, 2015, pp. 257).

3. Data and methods

In the following section provides an overview of the study design of the pilot study, the replication study, and the follow-up study.

3.1 Pilot study

The data used in this study was taken form Prolific (www.prolific.co/). To ensure high-quality work, the approval rate on Prolific was set to 95%. During the set-up of the study, it was ensured that only participants who self-identified as native English speakers based in the US were recruited. Every participant was allowed to participate only once and was paid for their participation. A total of 60 participants (35, females, 25 males) completed the study. Their ages range from 19 to 71 with a mean of 36.3. A total of 17 participants (%) completed high school as their highest level of education, while 27 participants (%) and 15 participants (%) reported college as their highest education or had a higher degree, respectively. A total of 29 (%)

participants described themselves as Democrats, 8 (13.3%) as Republicans, while 20 (33.3%) identified as neither and 2 (3.3%) preferred not to say.

The experiment was conducted using a 2 (metaphorical frame: *beast* or *virus*) x 2 (vignette length: short or long) between-subjects design. The four options of the metaphorically framed description of crime are given in the appendix.

After a welcoming page, participants were randomly presented with one of the four possible, metaphorically framed crime descriptions about crime in the City of Addison. While the description was still visible on the screen, participants were asked to propose solutions for the crime problem in a text box. Next, they were asked to rate the reliability of the text on a scale from 1 to 7. Finally, they were asked to indicate their political affiliation, age, education, gender, nationality, native language, and level of education. During the experiment, participants were not able to return to the pages they had already visited in the survey and update their responses.

Participants' solutions were treated as a set of suggestions. Each suggestion is classified as either *enforce* or *reform* by hand. As in the original study by Thibodeau & Boroditsky (2011), suggestions are categorized as *reform* if the proposed solution suggests investigating the underlying cause of the problem or suggests a particular social reform to treat or inoculate the community. In contrast, suggestions are categorized as *enforce* if the proposed solution focuses on the police force or other methods of law enforcement or modifying the criminal justice system. As pointed out by Thibodeau & Boroditsky (2015), suggestions of neighborhood watches cannot be clearly classified and are therefore excluded from the analysis as well. Solutions that contain the same number of suggestions for both reform and enforce are categorized as *both*. Finally, solutions are categorized as *neither* if the proposed solution lacked a suggestion and are therefore excluded from the analysis. The remaining three categories are ordered as follows: *reform*, *both*, *enforce*.

Data were analyzed using a Bayesian ordinal regression model for the binary choice variable `metaphor` as covariate (either *beast* or *virus*). The analysis uses the statistical programming language R and relies on the

`brms` package. In the analysis, the default (flat) priors of the `brms` package for the effect coefficients are used. The code for the analysis is given in the appendix.

3.2 Replication study

The data design was identical to the pilot study apart from the sample size. A total of 200 participants (92 females, 102 males, 3 other) completed the study. Their ages range from 18 to 1850 with a mean of 48. A total of 65 participants (32.5%) completed high school as their highest level of education, while 2 (1%) participants did not graduate from high school. 96 participants (%) reported college as their highest education, and 33 participants (16.5%) had a higher degree. A total of 108 (54%) participants described themselves as Democrats, 33 (16.5%) as Republicans, while 56 (28%) identified as neither and 3 participants (1.5%) preferred not to say.

The data analysis was mostly identical to the pilot study. However, because only 3 participants out of 200 participants preferred not to state their political affiliation, we excluded those participants from the analysis to conduct a reliable statistical analysis.

3.3 Follow-up study
To be inserted here

4. Findings

The aim of the replication study was to test whether metaphorically framing crime as a *beast* triggers participants to prefer enforcement-oriented solutions and describing crime as a *virus* influences participants to favor reform-oriented solutions, as has been claimed by Thibodeau & Boroditsky (2011). Additionally, the replication study explores the influence of the vignette length on the perceived reliability. Since the follow-up study tests the hypothesis that the effects of both the *beast* and the *virus* frame are higher if the speaker describing the crime issue seems reliable, it is crucial to ensure that it is not the vignette length but the perceived reliability of the speaker that amplifies such effects.

First consider Figure 1, in which the participants' responses to the crime issue are sorted into three ordered categories. It becomes apparent that enforcement-oriented solutions are preferred overall.
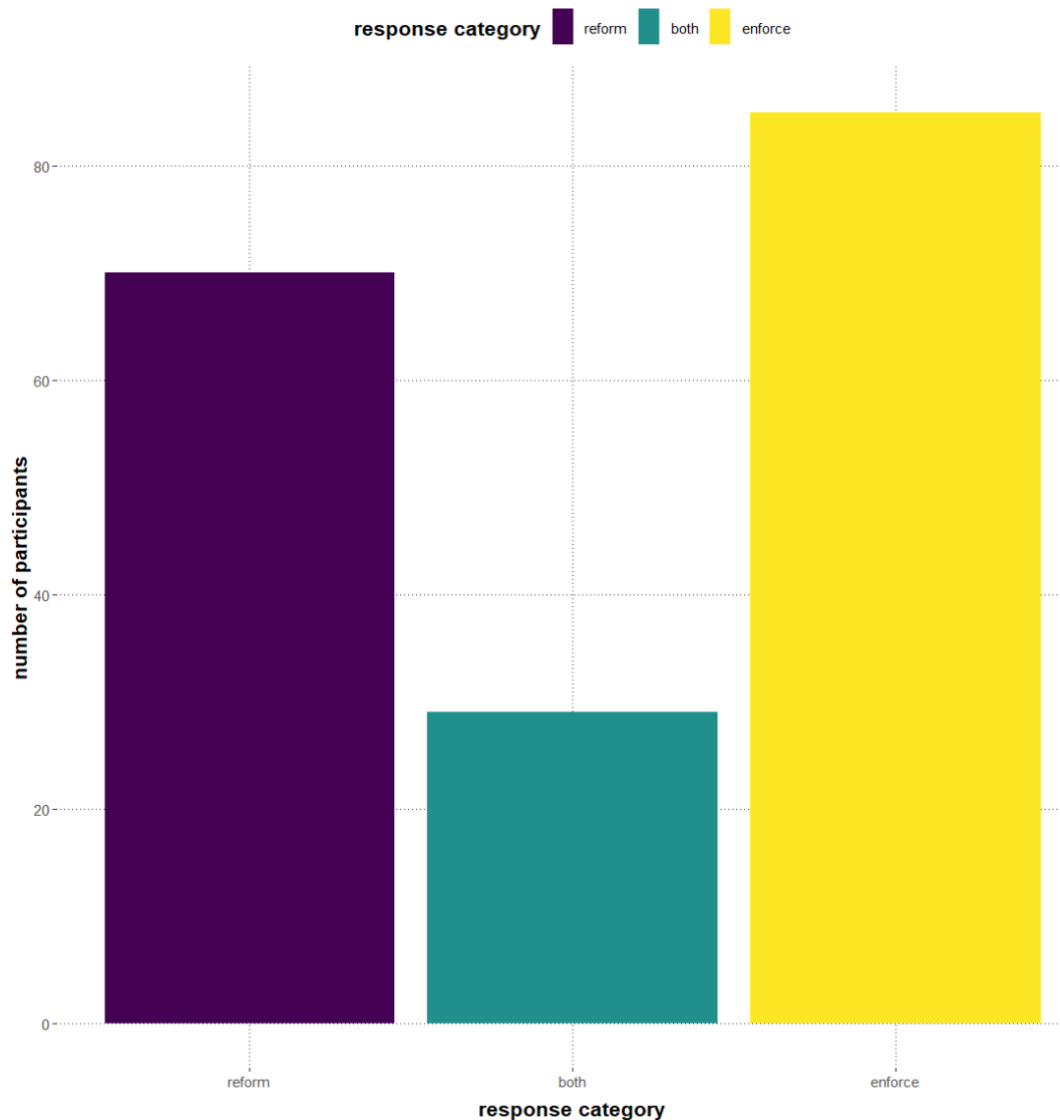

Figure 1: Categorizing participants' responses reveals that enforcement-oriented solutions are preferred overall.

Next consider Figure 2, which presents the same data but differentiates between political affiliations and the metaphorical framing of the description. Considered jointly, it becomes apparent that first, Republican mostly suggest enforcement-oriented solutions overall and second, metaphorically framing crime as a *virus* leads participants to favor reform-oriented solutions.
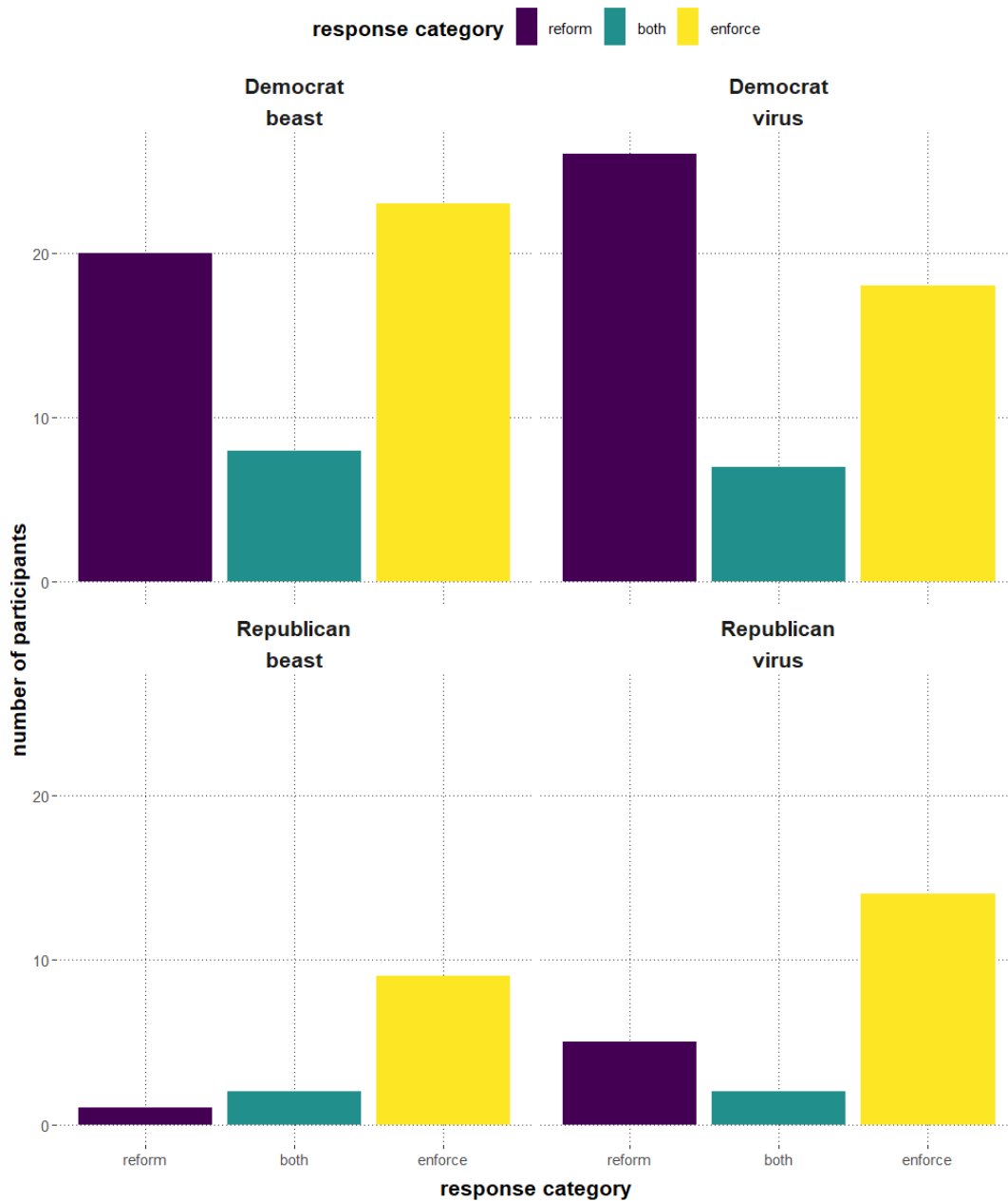
Figure 2: Participants' responses are influenced by the metaphorical framing of the description as well as their political affiliations.

The preference for enforcement-oriented measures after reading the metaphorical framing of crime as a *beast* rather than a *virus* fails to reach statistical significance ($p = 0.68$). However, the results of the present study could replicate the finding of the original study that Republicans, overall, are significantly more favor enforcement-oriented approaches than participants who identify their political affiliation as Democrats or as neither Republican nor Democrat ($p = 0.99$).

Lastly consider Figure 3, in which the influence of the vignette length on the perceived reliability is visualized. The mean values of the perceived reliabilities r are almost similar in both groups ($r = 4.8$ in the long vignette condition and $r = 5.1$ in the short vignette condition), although the results reveal that the variation is rather high ($\sigma = 1.5$ and $\sigma = 1.3$, respectively).
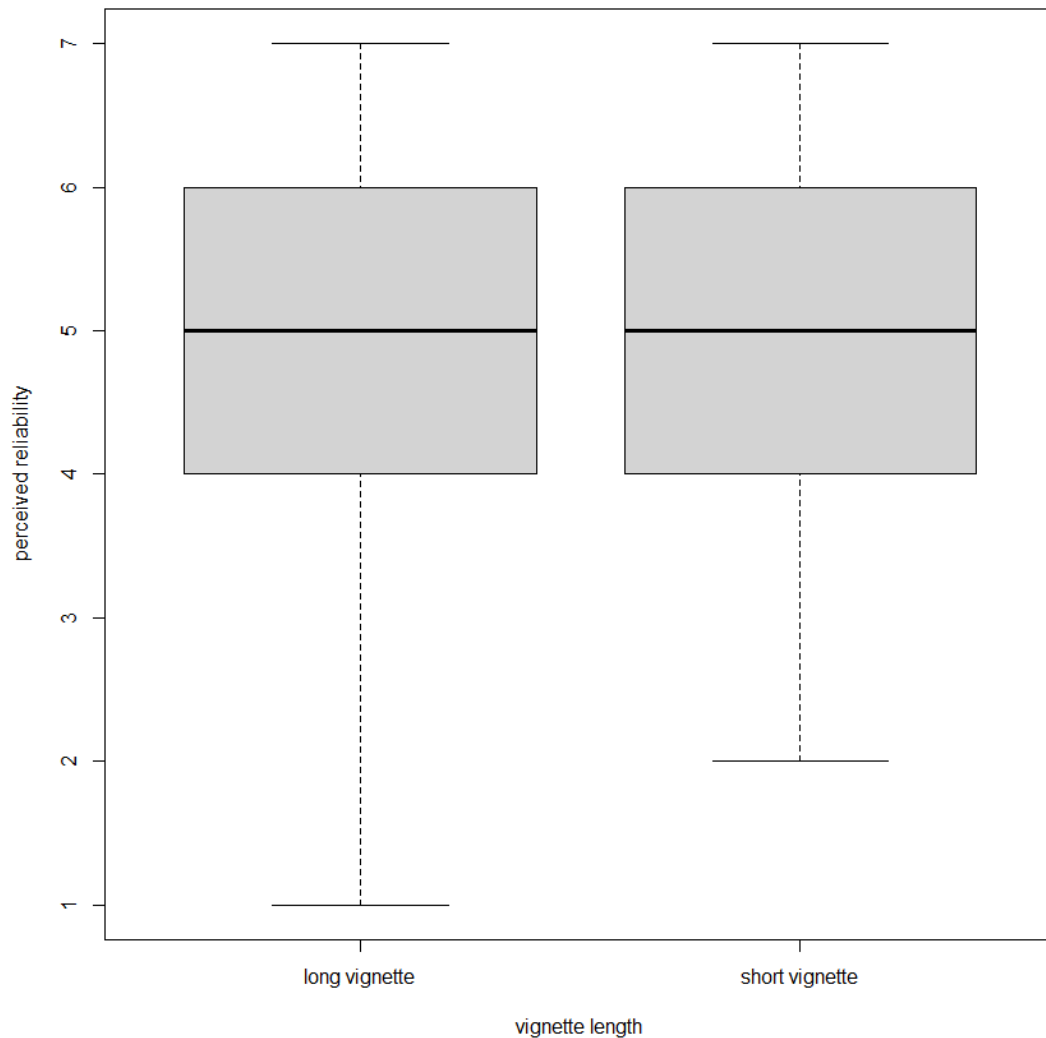

Figure 3: Plotting reliability against vignette length reveals hardly any influence of vignette length on perceived reliability.

As becomes apparent in the plot, the difference between both groups is not statistically significant ($p = 0.08$).

5. Discussion

To be inserted here

6. Conclusions

To be inserted here

7. Bibliography

Bougher, L. D. (2012). The Case for Metaphor in Political Reasoning and Cognition: Metaphor in Political Reasoning and Cognition. *Political Psychology*, *33*(1), 145–163. https://doi.org/10.1111/j.1467-9221.2011.00865.x

van den Broek, B. (2005). *Treating the virus and capturing the beast: The effect of verbal and visual metaphor on reasoning about social problems.* Master's thesis, Tilburg University.

Christmann, U., & Göhring, A.-L. (2016). A German-language replication study analysing the role of figurative speech in reasoning. *Scientific Data*, *3*(1), 160098. https://doi.org/10.1038/sdata.2016.98

Fausey, C. M. (2010). Constructing agency: The role of language. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00162

Flusberg, S. J., Lauria, M., Balko, S., & Thibodeau, P. H. (2020). Effects of Communication Modality and Speaker Identity on Metaphor Framing. *Metaphor and Symbol*, *35*(2), 136–152. https://doi.org/10.1080/10926488.2020.1767336

Flusberg, S. J., Matlock, T., & Thibodeau, P. H. (2018). War metaphors in public discourse. *Metaphor and Symbol*, *33*(1), 1–18. https://doi.org/10.1080/10926488.2018.1407992

Gibbs, R. W. (o. J.-a). *The Cambridge Handbook of Metaphor and Thought*.

Gibbs, R. W. (2015). Do pragmatic signals affect conventional metaphor understanding? A failed test of deliberate metaphor theory. *Journal of Pragmatics*, *90*, 77–87. https://doi.org/10.1016/j.pragma.2015.05.021

Heyvaert, P., Randour, F., Dodeigne, J., Perrez, J., & Reuchamps, M. (2020). Metaphors in political communication: A case study of the use of deliberate metaphors in non-institutional political interviews. *Journal of Language and Politics*, *19*(2), 201–225. https://doi.org/10.1075/jlp.17057.hey

Johannesson, N.-L. (o. J.). *When is metaphor deliberate*.

Knowles, M., & Moon, R. (2004). *Introducing Metaphor* (0 Aufl.). Routledge. https://doi.org/10.4324/9780203642368

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 585–589. https://doi.org/10.1016/S0022-5371(74)80011-3

Neuman, Y., Assaf, D., Cohen, Y., Last, M., Argamon, S., Howard, N., & Frieder, O. (2013). Metaphor Identification in Large Texts Corpora. *PLoS ONE*, *8*(4), e62343. https://doi.org/10.1371/journal.pone.0062343

Reijnierse, W. G., Burgers, C., Krennmayr, T., & Steen, G. J. (2015). How viruses and beasts affect our opinions (or not): The role of extendedness in metaphorical framing. *Metaphor and the Social World*, *5*(2), 245–263. https://doi.org/10.1075/msw.5.2.04rei

Steen, G. (2008). The Paradox of Metaphor: Why We Need a Three-Dimensional Model of Metaphor. *Metaphor and Symbol*, *23*(4), 213–241. https://doi.org/10.1080/10926480802426753

Steen, G. (2009). Deliberate Metaphor Affords Conscious Metaphorical Cognition. *Cognitive Semiotics*, *5*(1–2), 179–197. https://doi.org/doi:10.1515/cogsem.2013.5.12.179

Steen, G. (2017). Deliberate Metaphor Theory: Basic assumptions, main tenets, urgent issues. *Intercultural Pragmatics*, *14*(1), 1–24. https://doi.org/10.1515/ip-2017-0001

Steen, G. J. (2011). The contemporary theory of metaphor—Now new and improved! *Review of Cognitive Linguistics*, *9*(1), 26–64. https://doi.org/10.1075/rcl.9.1.03ste

Steen, G. J., Reijnierse, W. G., & Burgers, C. (2014). When Do Natural Language Metaphors Influence Reasoning? A Follow-Up Study to Thibodeau and Boroditsky (2013). *PLoS ONE*, *9*(12), e113536. https://doi.org/10.1371/journal.pone.0113536

Thibodeau, P. H. (2016). Extended Metaphors are the Home Runs of Persuasion: Don't Fumble the Phrase. *Metaphor and Symbol, 31*(2), 53–72. https://doi.org/10.1080/10926488.2016.1150756

Thibodeau, P. H., & Boroditsky, L. (2013). Natural Language Metaphors Covertly Influence Reasoning. *PLoS ONE, 8*(1), e52961. https://doi.org/10.1371/journal.pone.0052961

Thibodeau, P. H., & Boroditsky, L. (2015a). Measuring Effects of Metaphor in a Dynamic Opinion Landscape. *PLOS ONE, 10*(7), e0133939. https://doi.org/10.1371/journal.pone.0133939

Thibodeau, P. H., & Boroditsky, L. (2015b). Measuring Effects of Metaphor in a Dynamic Opinion Landscape. *PLOS ONE, 10*(7), e0133939. https://doi.org/10.1371/journal.pone.0133939

Thibodeau, P. H., Crow, L., & Flusberg, S. J. (2017). The metaphor police: A case study of the role of metaphor in explanation. *Psychonomic Bulletin & Review*, *24*(5), 1375–1386. https://doi.org/10.3758/s13423-016-1192-5

Thibodeau, P. H., & Durgin, F. H. (2011). Metaphor Aptness and Conventionality: A Processing Fluency Account. *Metaphor and Symbol*, *26*(3), 206–226. https://doi.org/10.1080/10926488.2011.583196

Thibodeau, P. H., & Flusberg, S. J. (In press). *Metaphor and Elaboration in Context*. John Benjamins Publishing. https://osf.io/c93nf/download

Thibodeau, P. H., Flusberg, S. J., Glick, J. J., & Sternberg, D. A. (2013a). An emergent approach to analogical inference. *Connection Science*, *25*(1), 27–53. https://doi.org/10.1080/09540091.2013.821458

Thibodeau, P. H., Flusberg, S. J., Glick, J. J., & Sternberg, D. A. (2013b). An emergent approach to analogical inference. *Connection Science*, *25*(1), 27–53. https://doi.org/10.1080/09540091.2013.821458

Thibodeau, P. H., Hendricks, R. K., & Boroditsky, L. (2017). How Linguistic Metaphor Scaffolds Reasoning. *Trends in Cognitive Sciences*, *21*(11), 852–863. https://doi.org/10.1016/j.tics.2017.07.001

Thibodeau, P. H., Matlock, T., & Flusberg, S. J. (2019). The role of metaphor in communication and thought. *Language and Linguistics Compass*, *13*(5), e12327. https://doi.org/10.1111/lnc3.12327

Thibodeau, P., Peebles, M. M., Grodner, D. J., & Durgin, F. H. (2015). The Wished-For Always Wins Until the Winner Was Inevitable All Along: Motivated Reasoning and Belief Bias Regulate Emotion During Elections: Elections Bias Reasoning. *Political Psychology*, *36*(4), 431–448. https://doi.org/10.1111/pops.12100

8. Appendix

8.1 Materials of replication study

Virus, long vignette:

Crime is a virus infecting on the city of Addison. The crime rate in the once peaceful city has steadily increased over the past three years. In fact, these days it seems that crime is plaguing every neighborhood. In 2004, 46,177 crimes were reported compared to more than 55,000 reported in 2007. The rise in violent crime is particularly alarming. In 2004, there were 330 murders in the city, in 2007, there were over 500.

Virus, short vignette:

Crime is a virus infecting on the city of Addison. The crime rate has steadily increased over the past three years. In fact, these days it seems that crime is plaguing every neighborhood. The rise in violent crime is particularly alarming.
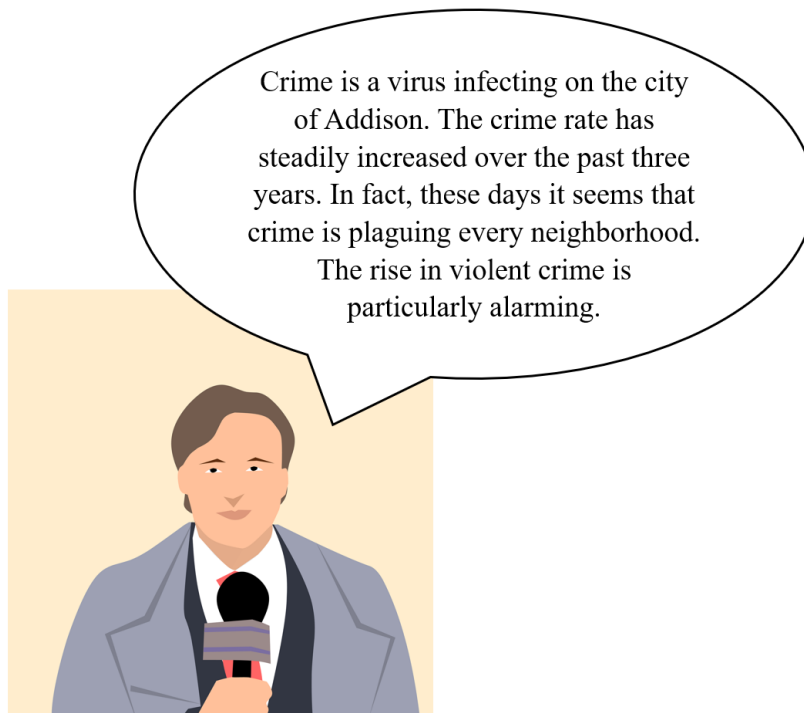
Beast, long vignette:

Crime is a wild beast preying on the city of Addison. The crime rate in the once peaceful city has steadily increased over the past three years. In fact, these days it seems that crime is lurking in every neighborhood. In 2004, 46,177 crimes were reported compared to more than 55,000 reported in 2007. The rise in violent crime is particularly alarming. In 2004, there were 330 murders in the city, in 2007, there were over 500.

Beast, short vignette:

Crime is a wild beast preying on the city of Addison. The crime rate has steadily increased over the last three years. In fact, these days it seems that crime is lurking in every neighborhood. The rise in violent crime is particularly alarming.

8.2 Materials of follow-up study

Virus, reliable speaker:

Crime is a virus infecting on the city of Addison. The crime rate has steadily increased over the past three years. In fact, these days it seems that crime is plaguing every neighborhood. The rise in violent crime is particularly alarming.

Virus, unreliable speaker:

Yo, crime is a damn virus infecting on our fucking city of Addison. The crime rate in the city went crazy in the last three years. Not gonna lie, now it seems that crime is plaguing every damn neighborhood. The rise in violent crime is fucking crazy.
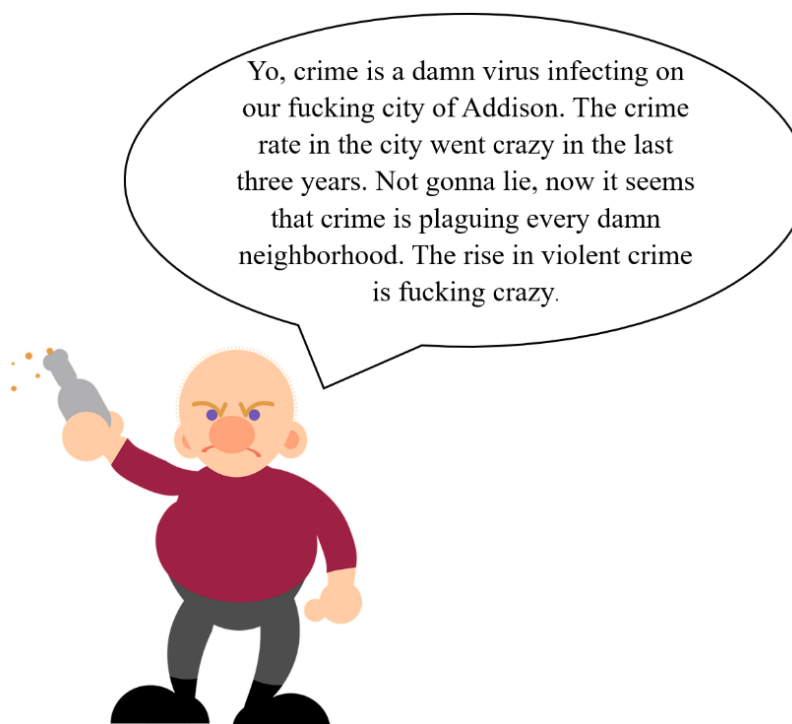
Beast, reliable speaker:

Crime is a wild beast preying on the city of Addison. The crime rate has steadily increased over the last three years. In fact, these days it seems that crime is lurking in every neighborhood. The rise in violent crime is particularly alarming.
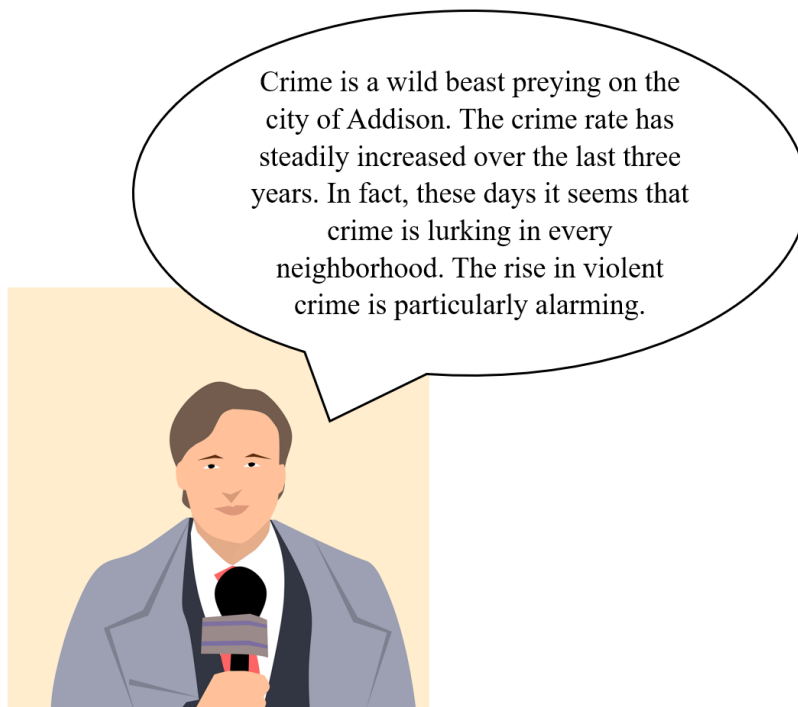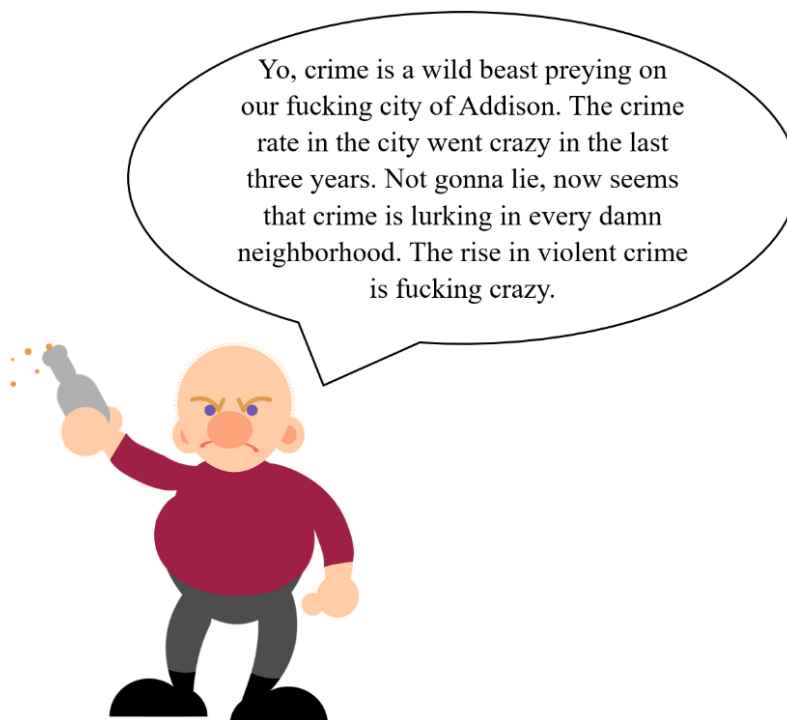
Beast, unreliable speaker:

Yo, crime is a wild beast preying on our fucking city of Addison. The crime rate in the city went crazy in the last three years. Not gonna lie, now seems that crime is lurking in every damn neighborhood. The rise in violent crime is fucking crazy.

8.3 Code for pilot study and replication study

```html
<template>
  <Experiment title="Study on problem solving">
    <InstructionScreen :title="'Welcome'">
      In this study, you are asked to come up with a method to solve
a problem.

      <br>

      Concretely, you will first have to read a background scenario.
      Please read it very carefully.
      Afterwards, you will be asked to come up with a potential so-
lution to the problem described in said scenario.
      Finally, you will be asked to rate the reliability of the text
describing the background scenario.
    </InstructionScreen>

    <Screen :validations="{
      response: {
        minLength: $magpie.v.minLength(5),
        required: $magpie.v.required
      }
    }">

      <Slide>
        <Record :data="{
            group: group,
            task: 'solution'
            }" />
        <p><div style='color:gray'>Background scenario</div></p>
        <p>{{background}}</p>
        <p><div style='color:gray'>Question</div></p>
        <p><strong>In your opinion what does Addison need to do to
reduce crime?</strong></p>
        <TextareaInput
          :response.sync= "$magpie.measurements.response"
        />
        <button v-if="!$magpie.validateMeasurements.response.$in-
valid" @click="$magpie.saveAndNextScreen();">Continue</button>
      </Slide>

    </Screen>


    <Screen :validations="{
      response: {
        required: $magpie.v.required
      }
```

```
    }">
      <Slide >
        <Record :data="{
              group: group,
              task: 'reliability'
              }" />
        Please rate the reliability of the text.
        <RatingInput quid="Quelle" :right="'very reliable'"
:left="'not reliable'" :response.sync="$magpie.measurements.re-
sponse" />

        <button v-if="!$magpie.validateMeasurements.response.$inva-
lid" @click="$magpie.saveAndNextScreen()">Submit</button>


      </Slide>
    </Screen>


    <Screen>

      <Slide>
        <Record :data="{
              group: group,
              task: 'affiliation'
              }" />
        <p>Please state your political affiliation.</p>
        <ForcedChoiceInput
            :response.sync= "$magpie.measurements.response"
            :options="['Democrat', 'Republican', 'neither', 'rather
not say']"
            @update:response="$magpie.saveAndNextScreen();"/>
      </Slide>

    </Screen>


    <PostTestScreen />
    <SubmitResultsScreen />
  </Experiment>
</template>

<script>
import _ from 'lodash';


// determine group

var group = _.shuffle(["beast, long vignette", "virus, long vi-
gnette", "beast, short vignette", "virus, short vignette"])[0]
```

```javascript
var backgrounds = {
    "beast, long vignette": "Crime is a wild beast preying on the
city of Addison. The crime rate in the once peaceful city has stead-
ily increased over the past three years. In fact, these days it
seems that crime is lurking in every neighborhood. In 2004, 46,177
crimes were reported compared to more than 55,000 reported in 2007.
The rise in violent crime is particularly alarming. In 2004, there
were 330 murders in the city, in 2007, there were over 500.",
    "virus, long vignette": "Crime is a virus infecting on the city
of Addison. The crime rate in the once peaceful city has steadily
increased over the past three years. In fact, these days it seems
that crime is plaguing every neighborhood. In 2004, 46,177 crimes
were reported compared to more than 55,000 reported in 2007. The
rise in violent crime is particularly alarming. In 2004, there were
330 murders in the city, in 2007, there were over 500.",
    "beast, short vignette": "Crime is a wild beast preying on the
city of Addison. The crime rate in the once peaceful city has stead-
ily increased over the past three years. In fact, these days it
seems that crime is lurking in every neighborhood. The rise in vio-
lent crime is particularly alarming.",
    "virus, short vignette": "Crime is a virus infecting on the city
of Addison. The crime rate in the once peaceful city has steadily
increased over the past three years. In fact, these days it seems
that crime is plaguing every neighborhood. The rise in violent crime
is particularly alarming.",
}

var background = backgrounds[group]

console.log(group)

export default {
  methods: {
    chooseQuestion: function () {
      return this.random;
    },
    randomOption: function () {
      const leftOption = this.options[Math.floor(Math.random() *
this.options.length)];
      const rightOption = this.options.find(option => option !==
leftOption);
      return [leftOption, rightOption]
    }
  },
  name: 'App',
  data() {
    return {
        group: group,
        background: background,
```

```
      options: [
        'Reform educational practices and create after school pro-
grams',
        'Increase street patrols that look for criminals'
      ],
      random: Math.random(),
      disableButton: true
    };
  },
  computed: {
    // Expose lodash to template code
    _() {
      return _;
    }
  }
};

</script>
```

8.4 Code for follow-up study

To be inserted here

8.5 Code for data analysis

```
---
title: "Data Analysis for Replication of the Role of Methaphor in Rea-
soning"
---

# set up
library("tidyverse")
library("brms")
library("faintr")
library("aida")

# these options help Stan run faster
options(mc.cores = parallel::detectCores())

#read in data of replication study
dat <- read_csv("results-replication-study.csv") %>%
  separate(
    # which column to split up
    col = group,
    # names of the new column to store results
    into = c("metaphor", "vignetteLength"),
    # separate by which character / reg-exp
    sep = ", ",
    # automatically (smart-)convert the type of the new cols
    convert = T
  )
```

```
dat <- dat %>%
  select(-responseTime) %>%
  pivot_wider(names_from = task, values_from = response)

# exclude data of pilot study
dat <- dat[61:260,]

# information about participants
# age
min(dat$age, na.rm=T)
max(dat$age, na.rm=T)
mean(dat$age, na.rm=T)

# gender
table(dat$gender)

# political affiliation
table(dat$affiliation)

# education
table(dat$education)

# data sorting by hand (see attached document "categorizing-of-re-
sponses.pdf")
# responses are categorized according to Thibodeau & Boroditsky (2011)
# 1=reform: proposed solution suggests investigating the underlying
cause of the problem or suggests a particular social reform to treat or
inoculate the community
# 2=enforce: proposed solution focuses on the police force or other
methods of law enforcement or modifying the criminal justice system
# 3=neither: proposed solution lacked a suggestion
# 4=both: proposed solution includes same number of suggestion for both
reform and enforce

# data sorting and cleaning
# solutions of the category "neither" are excluded according to
Thibodeau & Boroditsky (2011)
# suggestions in line of "neighborhood watches" are excluded according
to Thibodeau & Boroditsky (2015)

dat <- dat %>%
  mutate(
    #
    response_category = c("reform", "both", "both", "reform", "both",
"reform", "enforce", "enforce", "reform", "reform", "enforce", "en-
force", "reform", "reform", "reform", "enforce", "enforce", "reform",
"enforce", "reform", "neither", "reform", "enforce", "neither", "en-
force", "reform", "enforce", "enforce", "reform", "enforce", "enforce",
"both", "reform", "reform", "reform", "enforce", "both", "reform",
"both", "reform", "both", "reform", "reform", "reform", "reform",
"both", "both", "reform", "reform", "reform", "reform", "reform",
"both", "both", "reform", "enforce", "enforce", "reform", "neither",
"neither", "reform", "reform", "both", "both", "enforce", "enforce",
"reform", "enforce", "both", "reform", "enforce", "enforce", "reform",
"reform", "enforce", "reform", "reform", "reform", "reform", "neither",
"enforce", "both", "reform", "enforce", "enforce", "enforce", "enforce",
"neither", "enforce", "reform", "enforce", "reform", "reform", "en-
force", "enforce", "enforce", "reform", "enforce", "enforce", "enforce",
"reform", "reform", "enforce", "both", "reform", "enforce", "reform",
```

```
"both", "neither", "reform", "reform", "reform", "both", "enforce", "en-
force", "reform", "reform", "enforce", "both", "reform", "reform", "en-
force", "reform", "reform", "neither", "reform", "enforce", "reform",
"neither", "enforce", "reform", "enforce", "enforce", "enforce", "both",
"enforce", "both", "enforce", "enforce", "neither", "neither", "nei-
ther", "enforce", "enforce", "both", "both", "enforce", "reform", "en-
force", "enforce", "both", "enforce", "enforce", "enforce", "reform",
"enforce", "enforce", "both", "enforce", "neither", "enforce", "en-
force", "enforce", "reform", "enforce", "enforce", "reform", "enforce",
"both", "reform", "enforce", "enforce", "both", "enforce", "enforce",
"enforce", "enforce", "reform", "both", "enforce", "enforce", "enforce",
"enforce", "reform", "enforce", "enforce", "enforce", "enforce", "re-
form", "reform", "reform", "reform", "enforce", "enforce", "enforce",
"both", "reform", "enforce", "enforce", "enforce"),
    response_category = factor(response_category, ordered = T, levels =
c("reform", "both", "enforce", "neither"))
  )

nrow(dat)
# 200

# excluding participants whose proposed solutions did not include any
suggestions
dat <- dat[!dat$response_category == "neither",]
nrow(dat)
# 187

# excluding participants who chose "rather not say" for political affil-
iation to avoid issues associated with unequal sample sizes
dat <- dat[!dat$affiliation == "rather not say",]
nrow(dat)
# 184

# participants' comments on pilot study
dat %>% pull(comments) %>% unique()

# data plotting
dat %>%
  ggplot(aes(x = vignetteLength, y = as.numeric(reliability))) +
  geom_jitter(height = 0)

ggplot(data=dat, aes(x = affiliation, y = as.numeric(reliability), color
= vignetteLength)) +
  geom_jitter(height = 0)

boxplot(data=dat, as.numeric(reliability) ~ vignetteLength, xlab ="vi-
gnette length", ylab = "perceived reliability")

dat %>%
  ggplot(aes(x = response_category, fill = response_category, )) +
  geom_bar() + theme_aida() +
  labs(x = "response category", y = "number of participants", fill =
"response category")

dat %>%
  ggplot(aes(x = response_category, fill = response_category, )) +
  facet_wrap(affiliation ~ metaphor) +
  geom_bar() + theme_aida() +
```

```
  labs(x = "response category", y = "number of participants", fill =
"response category")

dat %>%
  ggplot(aes(x = response_category, fill = response_category, )) +
  facet_wrap(vignetteLength ~ metaphor) +
  geom_bar() + theme_aida()

# Hypotheses testing

# Hypothesis 1
# main effect of metaphor,
# specifically higher rates towards "enforce" for "beast" metaphor
# tested by comparing the posterior estimates for the aggregate value in
the "beast" condition with those from the "virus" condition
# as by this method:
fit <- brm(
  formula = response_category ~ metaphor * vignetteLength * affiliation,
  data = dat,
  family=cumulative("logit")
)

faintr::compare_groups(
  fit,
  higher = metaphor == "beast",
  lower  = metaphor == "virus"
)

# We judge there to be evidence in favor of the hypothesis, if the pos-
terior probability of this difference being bigger than zero is at least
0.95.
# results fail to reach significance

# We intend to also investigate the following two hypotheses

# Hypothesis 2
# Rates of "enforce" increase for participants that identify their po-
litical affiliation as "Republican".
# comparison with the grand mean
faintr::compare_groups(
  fit,
  higher = affiliation == "Republican"
)
# comparison with Democrats specifically
faintr::compare_groups(
  fit,
  higher = affiliation == "Republican",
  lower  = affiliation == "Democrat"
)

brms::hypothesis(fit, 'affiliationRepublican > 0')
brms::hypothesis(fit, 'metaphorvirus - affiliationRepublican > 0')

# We judge there to be evidence in favor of the hypothesis, if the pos-
terior probability of this difference being bigger than zero is at least
0.95.

# Hypothesis 3
```

```
# The effect of metaphors on reasoning, as proposed in hypothesis 1,
does not vary with the length of the metaphorically framed description.
dat %>%
  group_by(vignetteLength) %>%
  summarize(
    mean = mean(as.numeric(reliability)),
    SD = sd(as.numeric(reliability))
  )

fit2 <- brm(
  formula = as.numeric(reliability) ~ vignetteLength,
  data = dat
)

faintr::compare_groups(
  fit2,
  higher = vignetteLength == "long vignette",
  lower  = vignetteLength == "short vignette"
)

# We judge there to be evidence in favor of the hypothesis, if the pos-
terior probability of this difference not being bigger than zero is at
least 0.95.
```

Declaration of Authorship

I hereby confirm that this paper and the work presented in it is entirely my own. Where I have consulted the work of others this is always clearly stated. All statements taken literally from other writings or referred to by analogy are marked and the source is always given. This paper has not yet been submitted to another examination office, either in the same or similar form.

Tübingen, March 31st, 2023

Miriam Schiele

Miriam Schiele