

# Simulation Analyses

## Table of contents

|  |          |
|--|----------|
| <b>BRMS Simulation</b>   | <b>1</b> |
| Simulation method . . . . .  | 1        |
| Modelling . . . . .  | 2        |
| Number of participants best fit to the alternative model by the number of trials . . . | 2        |
| Number of participants with 95% credible estimates on the independence parameter       | 3        |
| Number of participants with 89% credible estimates on the independence parameter       | 5        |
| Looking at whether people had a preference for source diversity or repetition . . . .  | 5        |
| <b>Topic Analyses</b>  | <b>6</b> |
| Simulate some topic data (more topics) . . . . .                                       | 8        |

## BRMS Simulation

How many trials do we need to be able to detect an effect of source independence at the individual level?

### Simulation method

I simulated this data based on real participant data from Alister, Ransom, and Perfors (2022). There were two main manipulations in this study: 2 consensus direction (i.e., whether the members of the consensus were arguing for or against the claim) X 2 consensus independence (i.e., whether the members of the consensus all shared the same primary source, or all shared different sources). This means there were four “cells” in the experiment. For each cell, I calculated the mean and standard deviation belief update (the difference between their 1-100 belief in the claim before versus after seeing the consensus). Using these means and standard deviations, I simulated either 3, 6, 9, 12, 16, or 20 trials per cell. Because there were 4 cells, this resulted in 6 data sets with 12, 24, 36, 48, 64, or 80 trials per participant.

## Modelling

The goal of this simulation was to fit separate Bayesian linear models to each participant. Usually, what we do is fit these models to all of the participants, with participant as a random effect, in order to identify group level consensus effects, but this isn't able to capture individual differences.

I kept the modelling reasonably simple, with a null model that only considers the consensus direction:

```
post ~ prior + direction
```

Against an alternative model that also considers source independence:

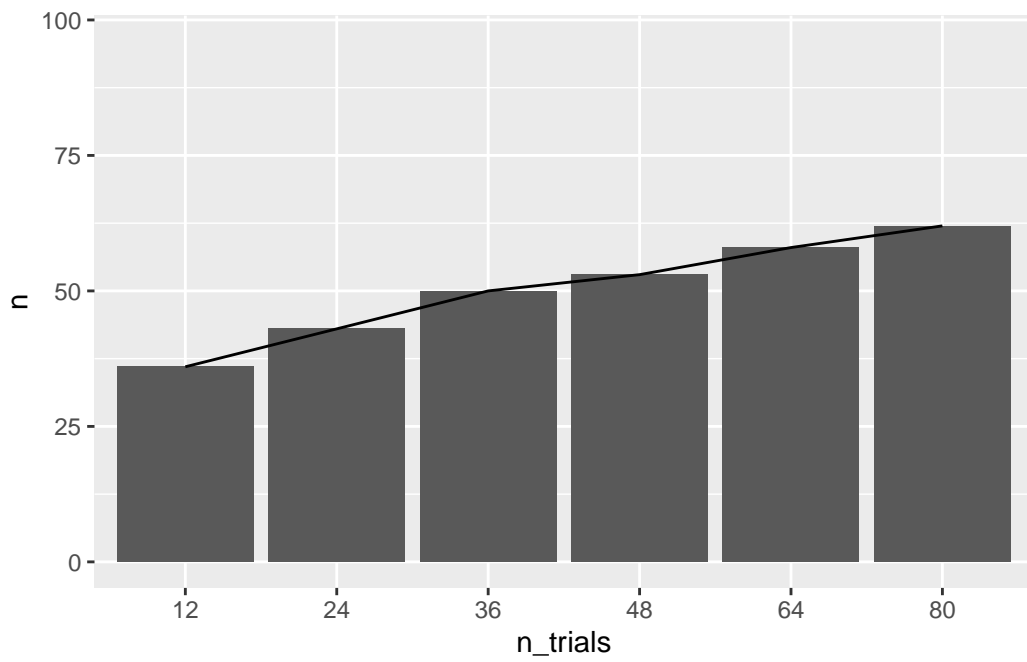
```
post ~ prior + direction + independence
```

One thing to keep in mind is that the models in the original study were not performed on the transformed "update" variable, but instead on their post score whilst accounting for their prior score as a control variable. I'm not sure how much of a difference this will make, but something to consider.

The general goal here is to see at which point increasing more trials does not appear to increase the number of participants best fit to the alternative model. Once that plateau begins to occur, then we should have a good idea of how many trials per participants we will need.

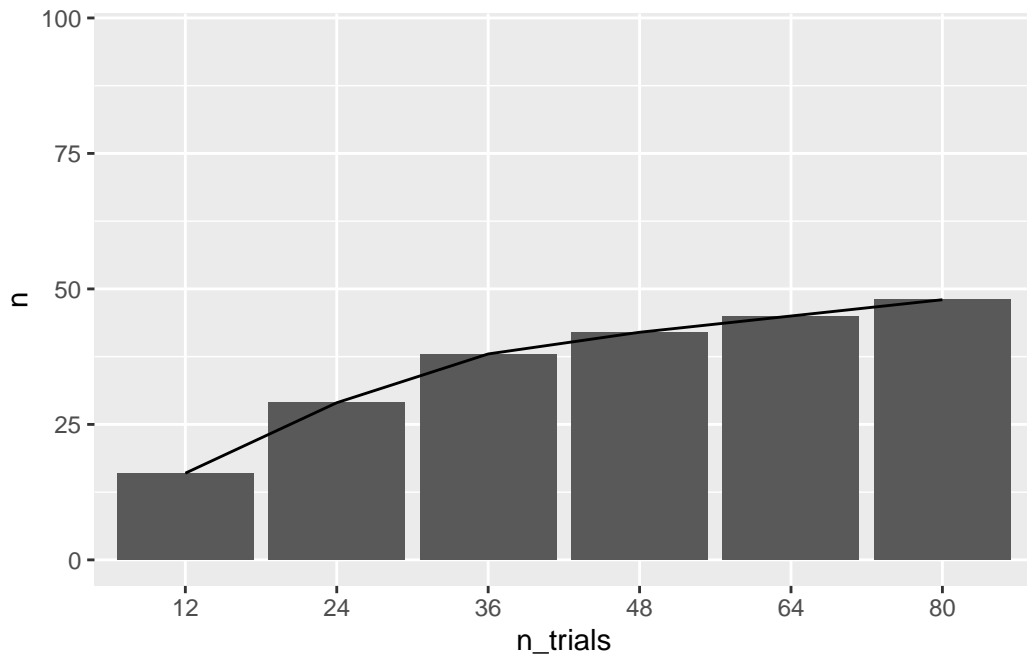
## Number of participants best fit to the alternative model by the number of trials

Using LOOIC, this analysis compared which participants were best fit to the alternative model versus the null model.



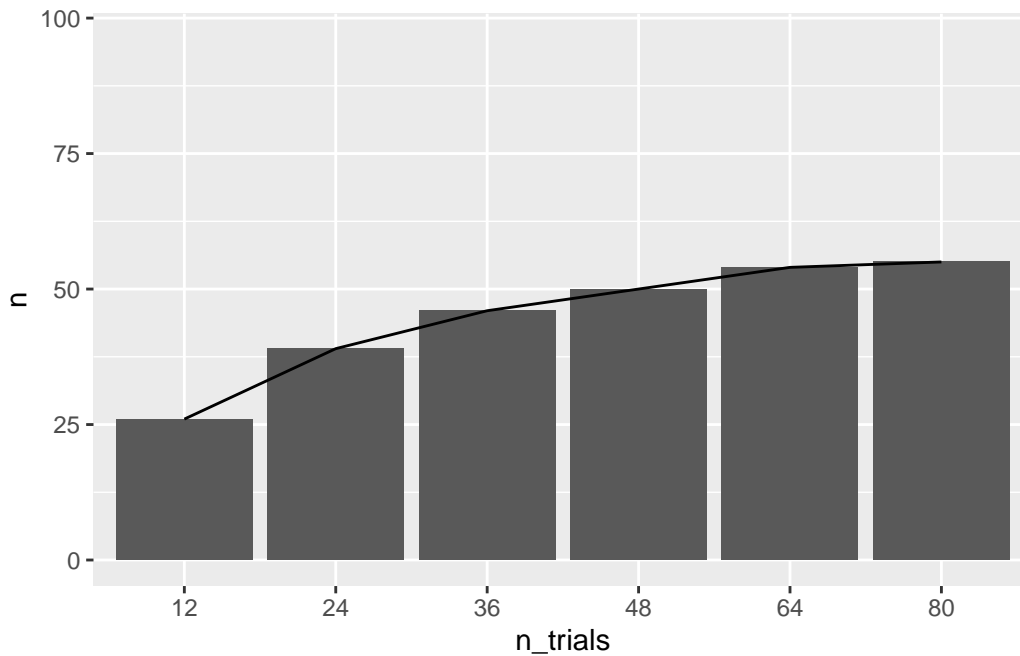
**Number of participants with 95% credible estimates on the independence parameter**

Another way to evaluate consensus independence is to look at the estimate and associated credible intervals of the independence parameter.



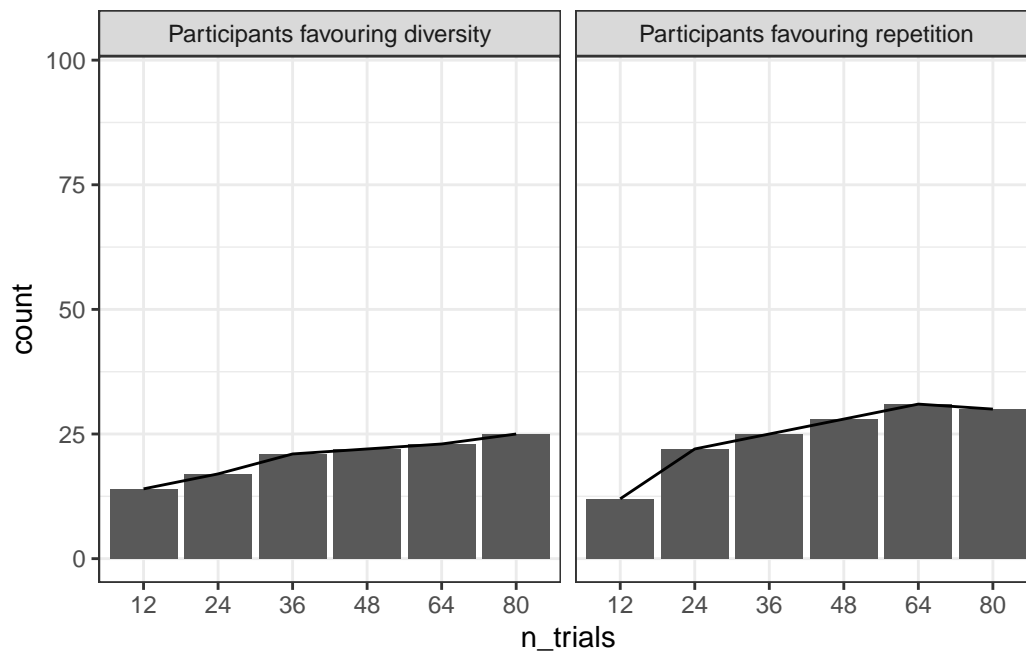
It has been argued that a 95% credible interval is not actually the most appropriate for Bayesian credible intervals, since unless you have an effective sample size (ESS) of at least 10,000, a 95% credible interval is not very stable (Kruschke 2014). An alternate interval is the 89% credible interval, which is much more stable for the ESS that we will have in our study (around 4000; Kruschke (2014)).

### Number of participants with 89% credible estimates on the independence parameter



### Looking at whether people had a preference for source diversity or repetition

A credible estimate of independence, or a model comparison favoring the alternative model only tells us whether there was an effect of consensus independence, but it does not tell us whether people preferred independent (diverse sources) or dependent (repeated) sources. People could reasonably have a preference for either, since diversity indicates more of the hypothesis spaces has been searched, whereas repetition might suggest that the source is more reliable, since it has been endorsed by more people. We can determine whether people had a preference for either by looking at the sign of the independence parameter. If it is positive, that person had a preference for diversity. If it was negative that person had a preference for repetition.



I'm not sure if this tells us much about experimental power, but interesting to see that in the simulated data at least, there different strategies are split pretty evenly. It is possible that the repetition effects are smaller, which is why there is a bigger difference in n as a function of trials.

## Topic Analyses

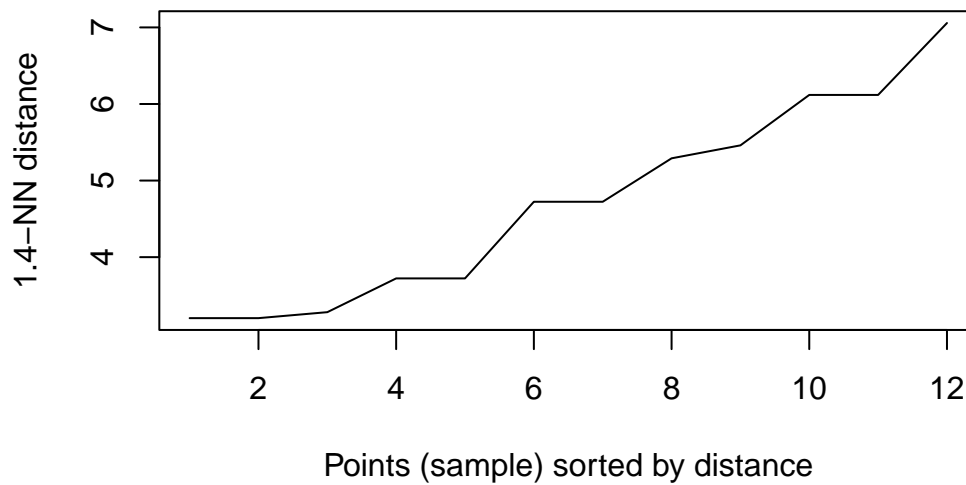
Are there some clusters which tend to elicit effects more than others?

One way— regression on each claim, get the estimates of those claims then cluster them?

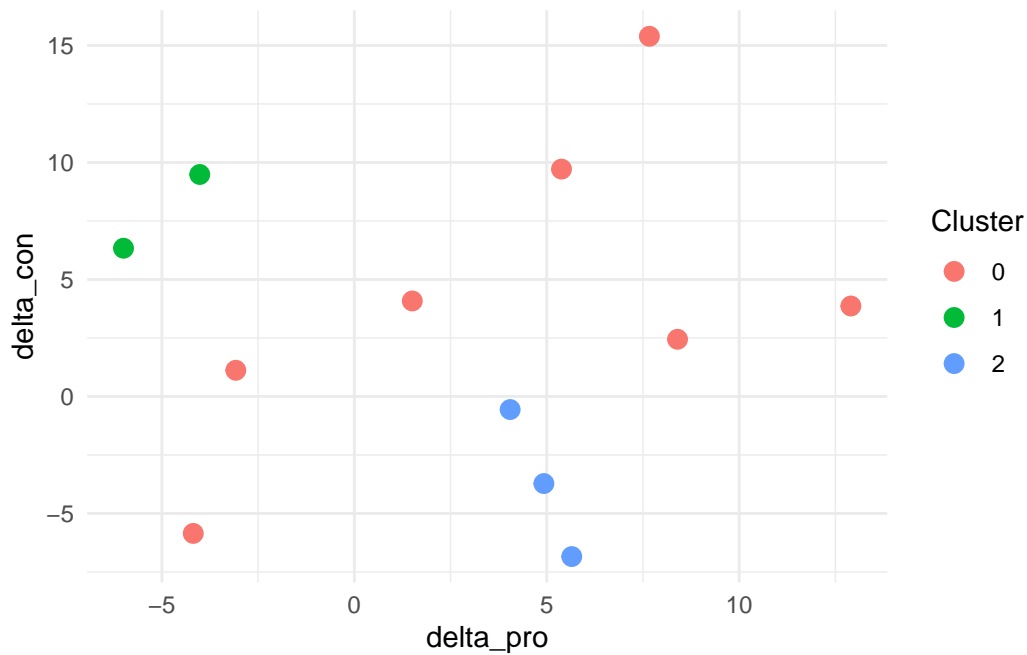
```
# A tibble: 12 x 3
# Groups:   trialType [12]
  trialType    delta_pro delta_con
  <fct>         <dbl>    <dbl>
1 fracking      -4.02     9.48
2 cleanCoal     -6.00     6.33
3 brexit         5.65    -6.84
4 xenophobia    -4.18    -5.85
5 charity       -3.08     1.12
6 lockdown       4.05    -0.560
7 avocado        8.40     2.44
8 narcissism     5.38     9.71
```

|    |               |      |       |
|----|---------------|------|-------|
| 9  | lockdownSleep | 7.67 | 15.4  |
| 10 | workFromHome  | 4.93 | -3.72 |
| 11 | writeOrType   | 1.51 | 4.08  |
| 12 | callOrText    | 12.9 | 3.86  |

```
# A tibble: 12 x 2
  delta_pro delta_con
  <dbl>      <dbl>
1    -4.02     9.48
2    -6.00     6.33
3     5.65    -6.84
4    -4.18    -5.85
5    -3.08     1.12
6     4.05    -0.560
7     8.40     2.44
8     5.38     9.71
9     7.67    15.4
10     4.93    -3.72
11     1.51     4.08
12    12.9     3.86
```



```
0 1 2
7 2 3
```



|             |            |            |               |              |
|-------------|------------|------------|---------------|--------------|
| fracking    | cleanCoal  | brexit     | xenophobia    | charity      |
| 1           | 1          | 2          | 0             | 0            |
| lockdown    | avocado    | narcissism | lockdownSleep | workFromHome |
| 2           | 0          | 0          | 0             | 2            |
| writeOrType | callOrText |            |               |              |
| 0           | 0          |            |               |              |

### Simulate some topic data (more topics)

With just 12 topics, this is probably very under powered. Let's simulate some more topics based on the actual mean and standard deviations of each topic according to each dimension and see how the topic estimates change.

Alister, Manikya, Keith James Ransom, and Andrew Perfors. 2022. "Source Independence Affects Argument Persuasiveness When the Relevance Is Clear." *Proceedings of the Annual Meeting of the Cognitive Science Society* 44 (44). <https://escholarship.org/uc/item/5hg4p8cm>.

Kruschke, John. 2014. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.