How Much Credible Interval Separation Rejects a Null Hypothesis?

Zhengxiao Wei [1], Farouk S. Nathoo [1], Michael E. J. Masson [2]
[1] Department of Mathematics and Statistics, [2] Department of Psychology, University of Victoria

Introduction

In null-hypothesis significance testing, two sample means are significantly different at the α level when the separation of their design-specific intervals is

$$|M_1 - M_2| \;/\; (2 \times \text{CI width}) > \sqrt{2}/2.$$

We examined five intervals in the analogous relationship between the Bayes factor and the interval separation.

The within-subject confidence interval, LM-CI.

The within-subject Bayesian interval, NKM-HDI, conditional on estimated random effects.

Two modifications of NKM-HDI, LH- and JZS-HDI, to allow for shrinkage and account for uncertainty in the estimation of random effects.

And, the standard highest-density interval.

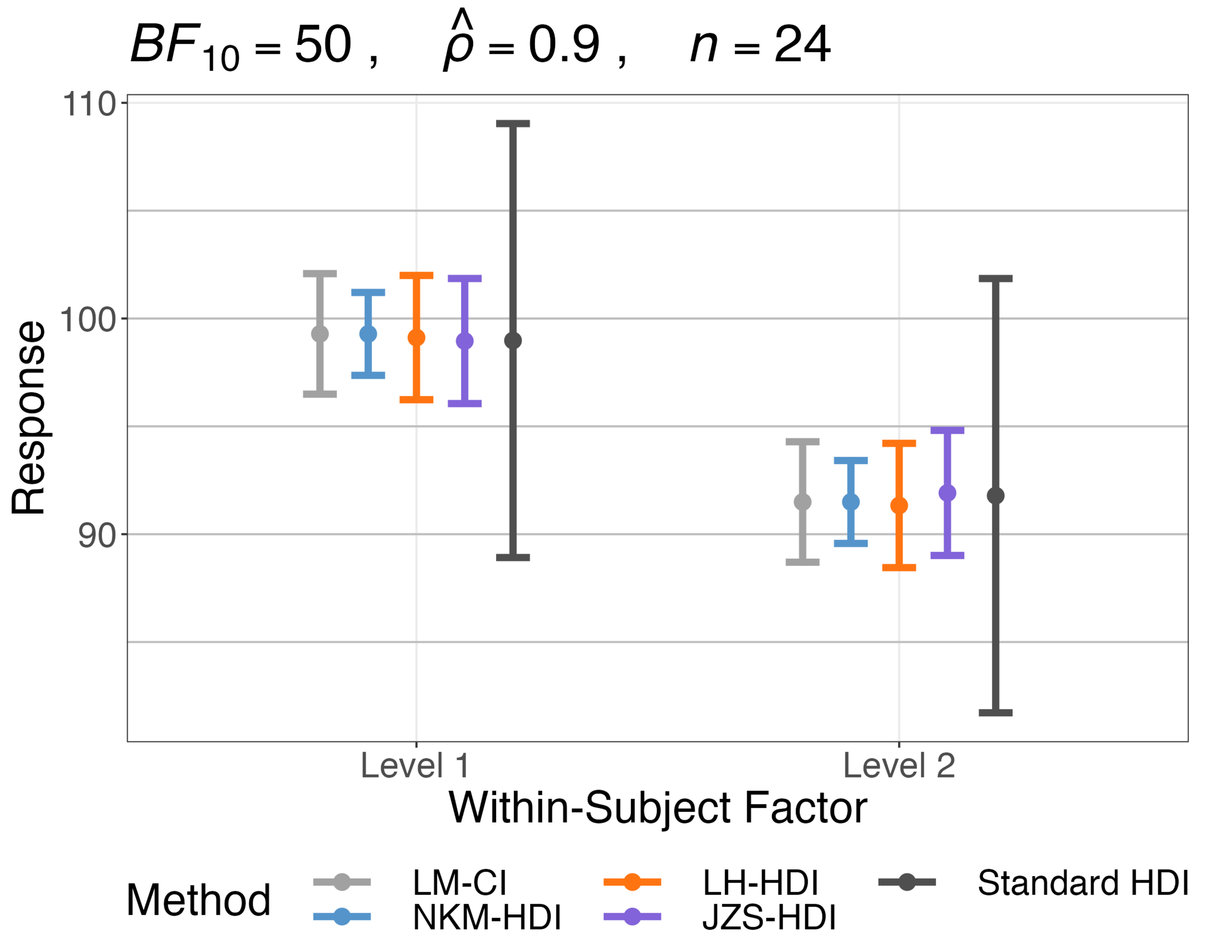

Figure 1: Five interval estimates for the condition means and very strong evidence supporting an effect from the Bayes factor in a one-way within-subject design.

Standard HDI is relatively wide due to the nuisance between-subjects variability and tends to obscure effects when used in within-subject designs

Benchmarking

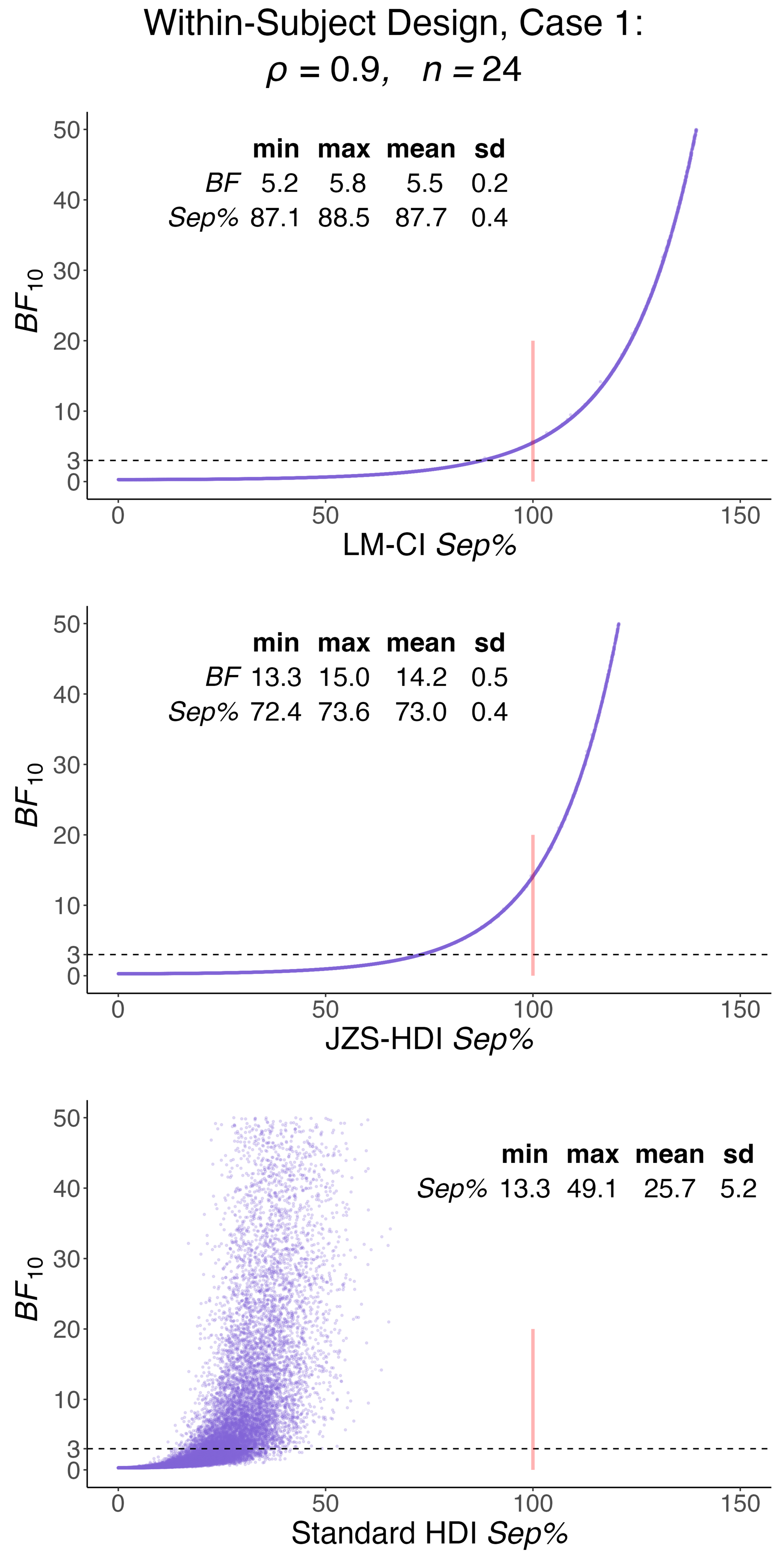

Figure 2: Scatter plots of the relationship in a selected method and benchmarks: the Bayes factor when interval separation is 100%; the separation percentage when the Bayes factor is 3 (moderate evidence)

Importance of Sample Size and Evidence for the Null

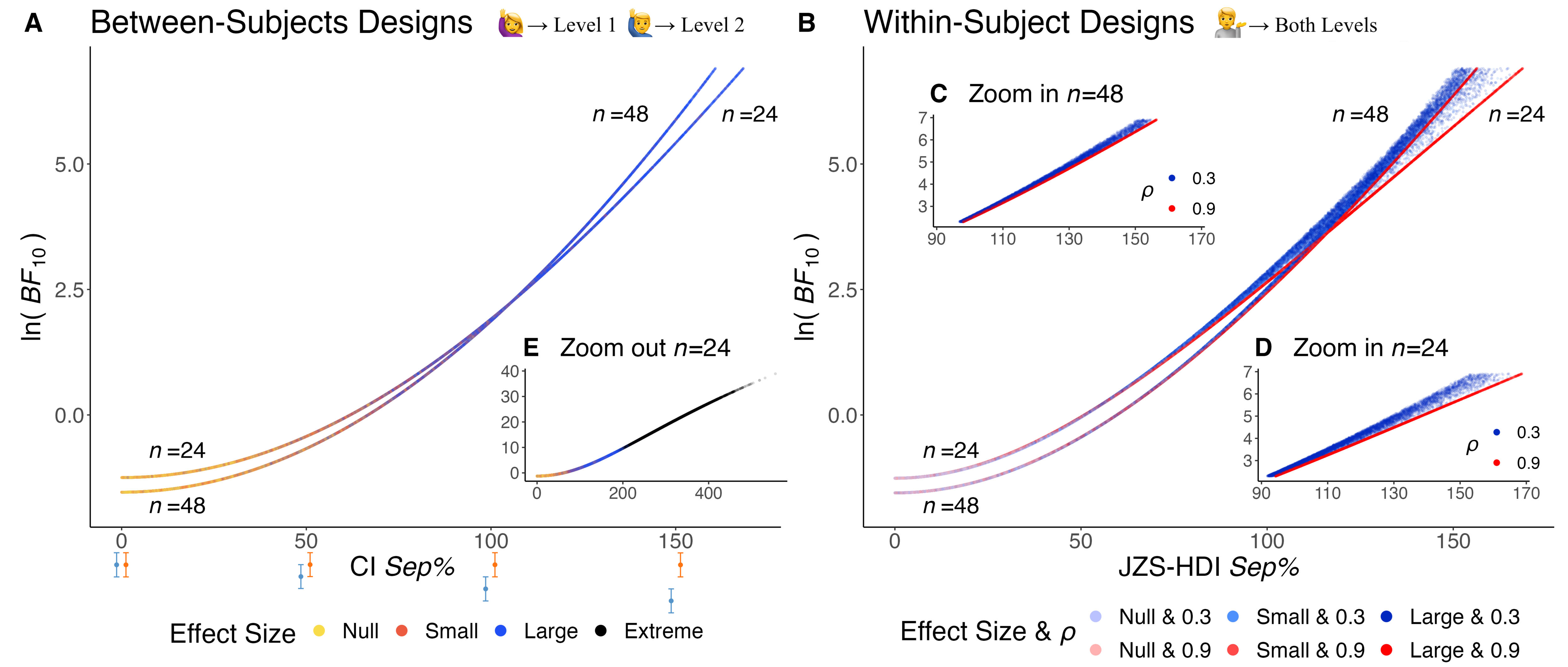

Figure 3: Effect sizes constitute different sections of the curve, dependent on sample size, in both designs. The greater the correlation in a within-subject design, the more consistent the observed quadratic exponential relationship is

Methods

$$\text{Standard CI} = M_{i\cdot} \pm \sqrt{\frac{SS_W}{n(n-1)a}} \cdot t^*_{1-\frac{\alpha}{2},\, a(n-1)}$$

$$\text{LM-CI} = M_{i\cdot} \pm \sqrt{\frac{SS_{S\times C}}{n(n-1)(a-1)}} \cdot t^*_{1-\frac{\alpha}{2},\, (n-1)(a-1)}$$

$$\text{NKM-HDI} = M_{i\cdot} \pm \sqrt{\frac{SS_{S\times C}}{n(n-1)a}} \cdot t^*_{1-\frac{\alpha}{2},\, a(n-1)}$$

$$\text{LH- or JZS-HDI} = \mathbb{E}\left[\mu_i \pm \frac{\sigma_\epsilon}{\sqrt{n}} \cdot t^*_{1-\frac{\alpha}{2},\, a(n-1)} \;\middle|\; \boldsymbol{y}\right]$$

- Separation :=
  The distance between two interval centers divided by the root mean square of interval lengths.

$$BF_{10} = \frac{\int_{\boldsymbol{\theta}\in\Theta_1} p(\boldsymbol{y} \mid \boldsymbol{\theta}, \mathcal{M}_1) \cdot p(\boldsymbol{\theta} \mid \mathcal{M}_1)\, d\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in\Theta_0} p(\boldsymbol{y} \mid \boldsymbol{\theta}, \mathcal{M}_0) \cdot p(\boldsymbol{\theta} \mid \mathcal{M}_0)\, d\boldsymbol{\theta}}$$

- The Bayes factor provides graded evidence of how data supports one model over another.

Concluding Remarks

We observed through simulations

a quadratic exponential relationship between the Bayes factor and the separation of credible intervals

in the linear model for a between-subjects design and the linear mixed-effects model for a within-subject design, each with two levels and a known sample size.

We provide a Stan-based R package '*rmBayes*' to enable computation of each of the within-subject credible intervals investigated here using a number of possible priors.

```
install.packages("rstan")
install.packages("rmBayes", type="binary")
```

Reference

Wei, Z., Nathoo, F. S., & Masson, M. E. J. (2023). Investigating the relationship between the Bayes factor and the separation of credible intervals. *Psychon. Bull. Rev.*, *30*, 1759–1781. doi: 10.3758/s13423-023-02295-1

Acknowledgements

This work was supported by discovery grants to Farouk S. Nathoo and Michael E. J. Masson from the Natural Sciences and Engineering Research Council.