1. Anagram task

Basic

"We defined two models to describe our data: $\mathcal{H}_1$ states that mean solution time for participants in the organized list condition will be less than the mean solution time for participants in the unrelated list condition, whereas $\mathcal{H}_0$ states that mean solution time will not differ between conditions. We then computed a Bayesian independent samples t-test (Rouder et al., 2009) to quantify the evidence for $\mathcal{H}_1$ over $\mathcal{H}_0$. We found a Bayes factor of $B_{10} = 7.11$, indicating that the observed data are approximately 7 times more likely under $\mathcal{H}_1$ than $\mathcal{H}_0$."

Advanced

"We defined two models to describe our data: $\mathcal{H}_1 : \delta < 0$ states that mean solution time for participants in the organized list condition will be less than the mean solution time for participants in the unrelated list condition, whereas $\mathcal{H}_0 : \delta = 0$ states that mean solution time will not differ between conditions. We then computed a Bayesian independent samples t-test (Rouder et al., 2009) to quantify the evidence for $\mathcal{H}_1$ over $\mathcal{H}_0$. This test requires the user to specify a prior distribution for effect size δ, which we initially took at the default Cauchy prior with scale $r = 0.707$. Using this prior, we found a Bayes factor of $B_{10} = 7.11$, indicating that the observed data are approximately 7 times more likely under $\mathcal{H}_1$ than $\mathcal{H}_0$. Additionally, we performed a robustness check by varying the prior scale factor r, each reflecting a different *a priori* expectation of the effect of our manipulation. Generally, B_{10} decreases as the scale factor r increases, but even using a very wide prior with $r = 1.41$, the inference remains the same, as the data are approximately 4.5 times more likely under $\mathcal{H}_1$ than under $\mathcal{H}_0$. "

2. Kitchen rolls experiment

Basic

"We defined two models to describe our data: $\mathcal{H}_1$ states that mean NEO score for participants in the clockwise condition will be greater than the mean NEO score for participants in the counterclockwise condition, whereas $\mathcal{H}_0$ states that mean NEO score will not differ between conditions. We then computed a Bayesian independent samples t-test (Rouder et al., 2009) to quantify the evidence for $\mathcal{H}_0$ over $\mathcal{H}_1$. We found a Bayes factor of $B_{01} = 7.80$, indicating that the observed data are approximately 7 times more likely under $\mathcal{H}_0$ than $\mathcal{H}_1$."

Advanced

"We defined two models to describe our data: $\mathcal{H}_1 : \delta > 0$ states that mean NEO score for participants in the clockwise condition will be greater than the mean NEO score for participants in the counterclockwise condition, whereas $\mathcal{H}_0 : \delta = 0$ states that mean NEO score will not differ between conditions. We then computed a Bayesian independent samples t-test (Rouder et al., 2009) to quantify the evidence for $\mathcal{H}_0$ over $\mathcal{H}_1$. This test requires the user to specify a prior distribution for effect size δ, which we initially took at the default Cauchy prior with scale $r = 0.707$. Using this prior, we found a Bayes factor of $B_{01} = 7.80$, indicating that the observed data are approximately 7 times more likely under $\mathcal{H}_0$ than $\mathcal{H}_1$. Additionally, we performed a robustness check by varying the prior scale factor r, each reflecting a different *a priori* expectation of the effect of our manipulation. B_{01} increases as the scale factor r increases ($BF_{01} = 10.8$ with scale $r = 1$ and $BF_{01} = 15.1$ with scale $r = 1.41$). Thus, our data is strongly evidential of $\mathcal{H}_0$ over $\mathcal{H}_1$."

3. Recall task

" We defined five competing models for our 2x2 factorial design:

- $\mathcal{M}_0$: null model
- $\mathcal{M}_1$: encodingCue only
- $\mathcal{M}_2$: retrievalCue only
- $\mathcal{M}_3$: encodingCue + retrievalCue
- $\mathcal{M}_4$: encodingCue + retrievalCue + encodingCue*retrievalCue

These models were compared via a Bayesian analysis of variance (Rouder et al., 2012). Equal prior probabilities were assigned to the five competing models. The additive model received the largest posterior probability, $p(\mathcal{M}_4 \mid \text{data}) = 0.621$. This was larger than the posterior probability of the interactive model, which received a posterior probability of $p(\mathcal{M}_4 \mid \text{data}) = 0.156$, as well as the posterior probability of the retrieval cue only model, $p(\mathcal{M}_2 \mid \text{data}) = 0.179$. In terms of model odds, the additive model received the most support from the data, which updated the prior odds for the model by a factor of 6.56. In comparison, the prior odds for all other models were actually decreased by the data. Finally, the Bayes factor directly comparing $\mathcal{M}_3$ to $\mathcal{M}_4$ was equal to 3.99, indicating that the observed data is approximately 4 times more likely under the additive model than under the interactive model."

4. Mental arithmetic task

" We defined five competing models for our 2x2 factorial design:

- $\mathcal{M}_0$: null model
- $\mathcal{M}_1$: format only
- $\mathcal{M}_2$: problem size only
- $\mathcal{M}_3$: format + problem size
- $\mathcal{M}_4$: format + problem size + format*problem size

These models were compared via a Bayesian analysis of variance (Rouder et al., 2012). Equal prior probabilities were assigned to the five competing models. The interactive model received the largest posterior probability, $p(\mathcal{M}_4 \mid \text{data}) = 0.998$. This was much larger than the posterior probability of the additive model, which received a posterior probability of $p(\mathcal{M}_3 \mid \text{data}) = 0.002$. In terms of model odds, the interactive model received the most support from the data, which updated the prior odds for the model by a factor of 2,385. In comparison, the prior odds for the additive model were actually decreased by a factor of $1/0.007 = 143$. Finally, the Bayes factor directly comparing $\mathcal{M}_4$ to $\mathcal{M}_3$ was equal to 596.2, indicating that the observed data is approximately 600 times more likely under the interactive model than under the additive model."

References:

1. Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review, 16*, 225-237.

2. Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374.