UNIVERSITY OF AMSTERDAM

Does Everyone? Modeling Individual Differences in Cognitive Tasks.

Julia Haaf

September, 2019

Psychology and Statistics

- Get use new insights using new scientific methods to advance psychology.

Psychology and Statistics

- Get use new insights using new scientific methods to advance psychology.
- Found scientific methods in theoretical considerations.

Psychology and Statistics

- Get use new insights using new scientific methods to advance psychology.
- Found scientific methods in theoretical considerations.
- When combining methods and theory we can come up with new and interesting research questions and solutions.

Theoretical Statements

- Levels of processing (Craik & Lockhart, 1972): *Processing words semantically increases memory performance as compared to more shallow processing.*

Psychology Guiding Methods

Theoretical Statements

- Levels of processing (Craik & Lockhart, 1972): *Processing words semantically increases memory performance as compared to more shallow processing.*
- Stroop interference (Stroop, 1935): *Due to interference, processing of incongruent items takes longer than processing of congruent items.*

The Stroop Effect

ROT ROT

Psychology Guiding Methods

Theoretical Statements

- Levels of processing (Craik & Lockhart, 1972): *Processing words semantically increases memory performance as compared to more shallow processing.*
- Stroop interference (Stroop, 1935): *Due to interference, processing of incongruent items takes longer than processing of congruent items.*

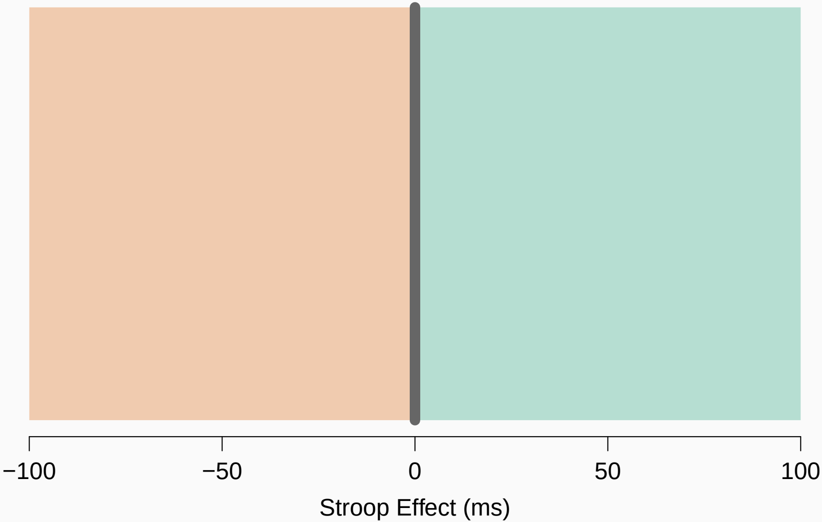
Ordinal Constraints in Social Sciences

- Claim: Most psychological theory makes ordinal predictions on data.

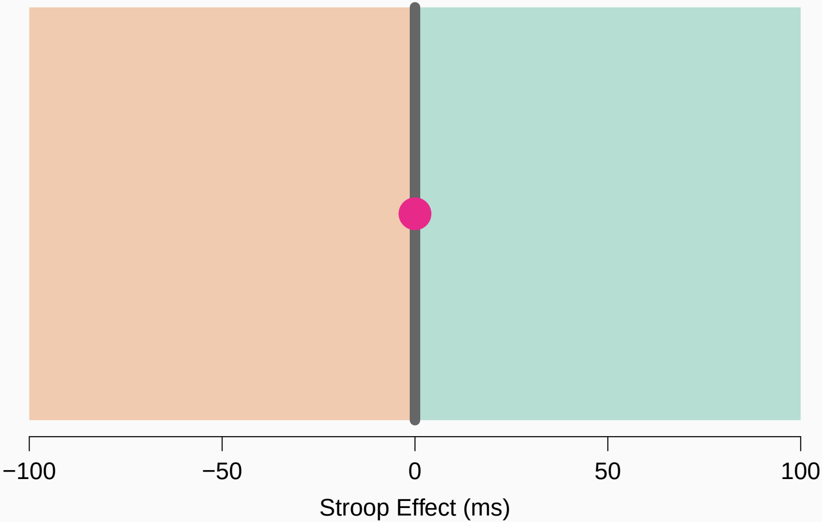
Ordinal Constraints in Social Sciences

- Claim: Most psychological theory makes ordinal predictions on data.
- Almost never metric predictions.

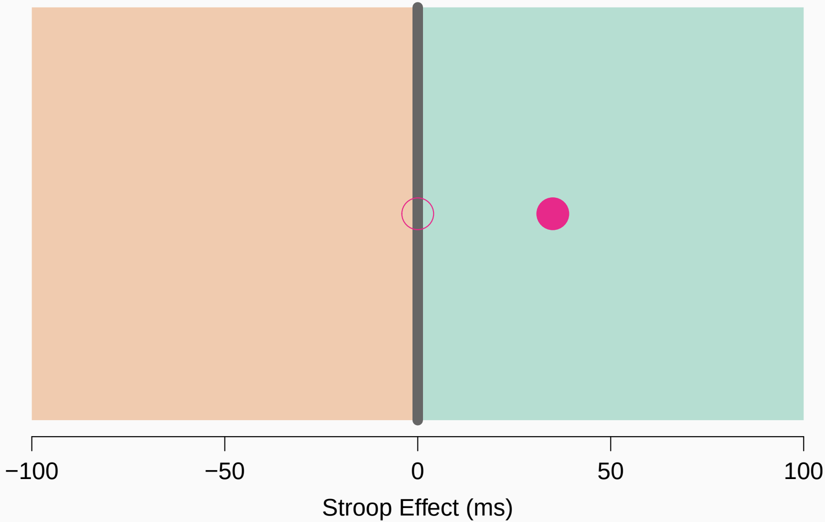
Ordinal Constraints and the Stroop Effect



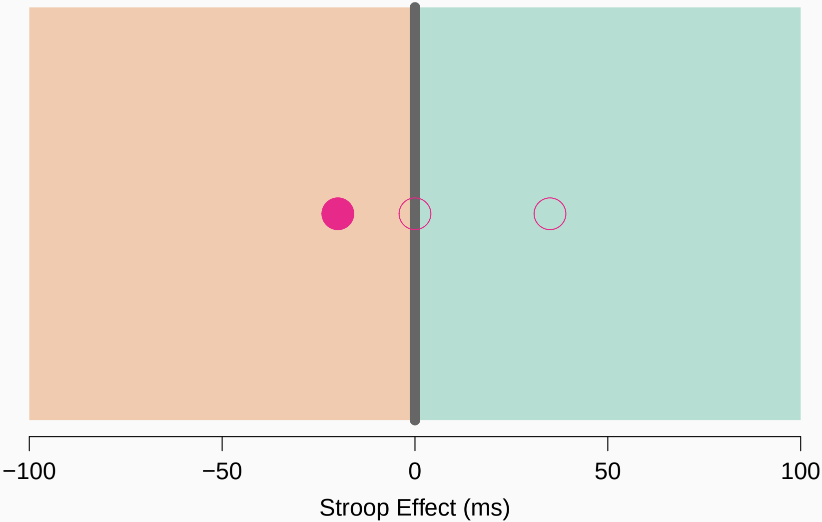
Ordinal Constraints and the Stroop Effect



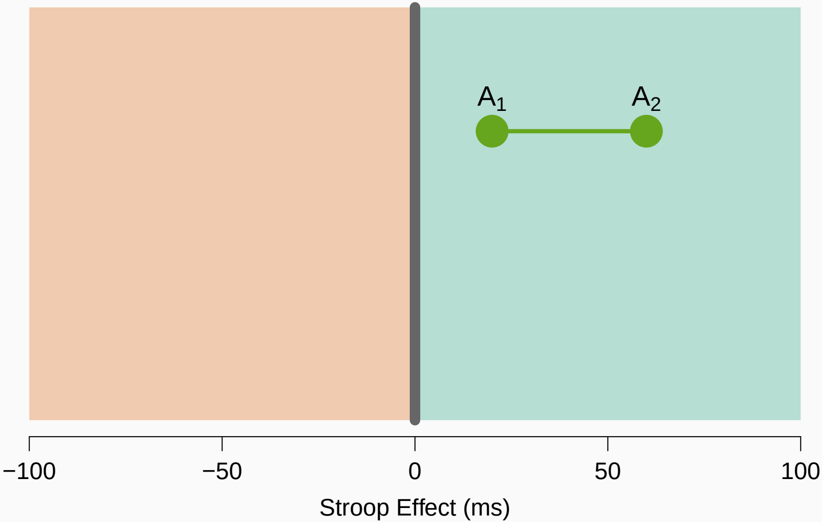
Ordinal Constraints and the Stroop Effect



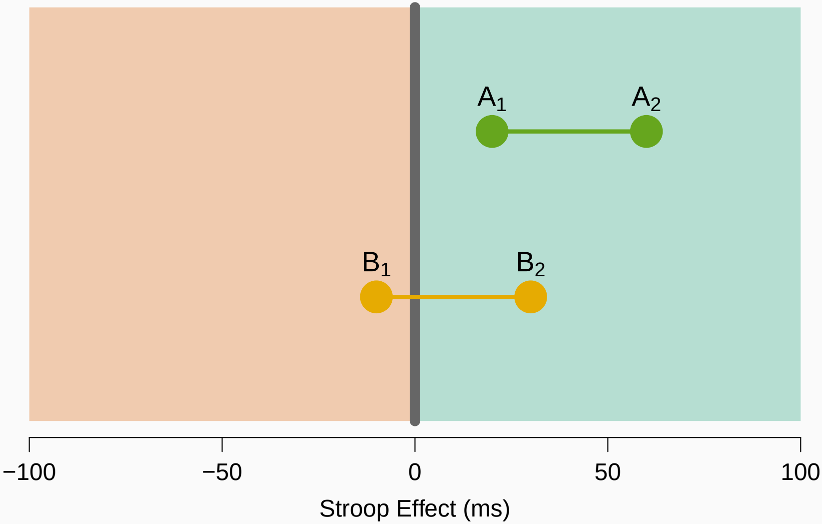
Ordinal Constraints and the Stroop Effect



Quantitative and Qualitative Individual Differences



Quantitative and Qualitative Individual Differences



Psychology Guiding Methods

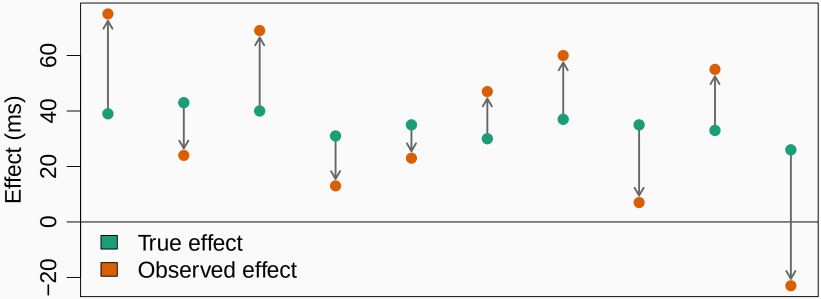
- Statistical models of Stroop interference should respect the theoretically important areas of negative, null, and positive effects.

Psychology Guiding Methods

- Statistical models of Stroop interference should respect the theoretically important areas of negative, null, and positive effects.
- Statistical models should distinguish between qualitative and quantitative individual differences.

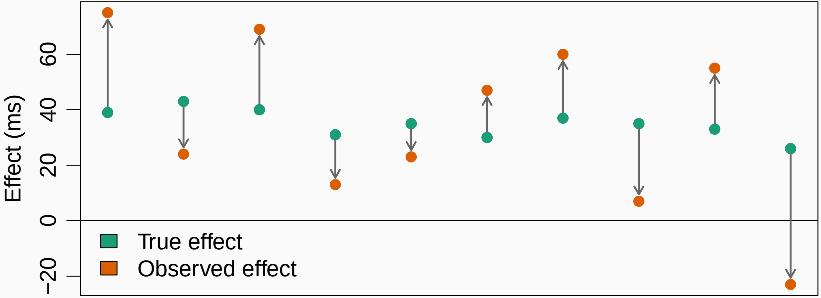
Methods Guiding Psychology

- Before modeling we need to consider the effect of sample noise in this setting.



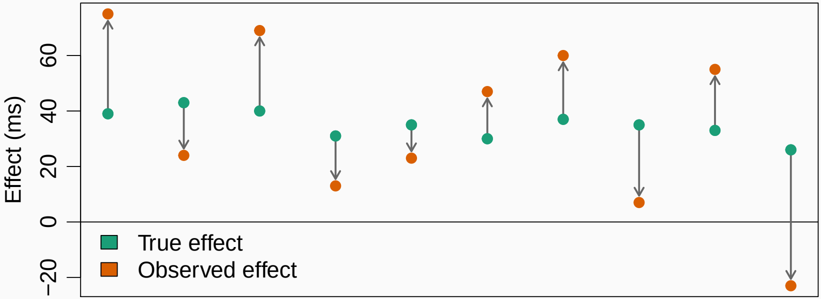
Methods Guiding Psychology

- Before modeling we need to consider the effect of sample noise in this setting.
- *Observed effect = True effect + sample noise* $\rightarrow$ What we observe with limited numbers of trials.



Methods Guiding Psychology

- Before modeling we need to consider the effect of sample noise in this setting.
- *Observed effect = True effect + sample noise* → What we observe with limited numbers of trials.
- *True effect*: What we would obtain if we had an unlimited number of trials per person per condition.

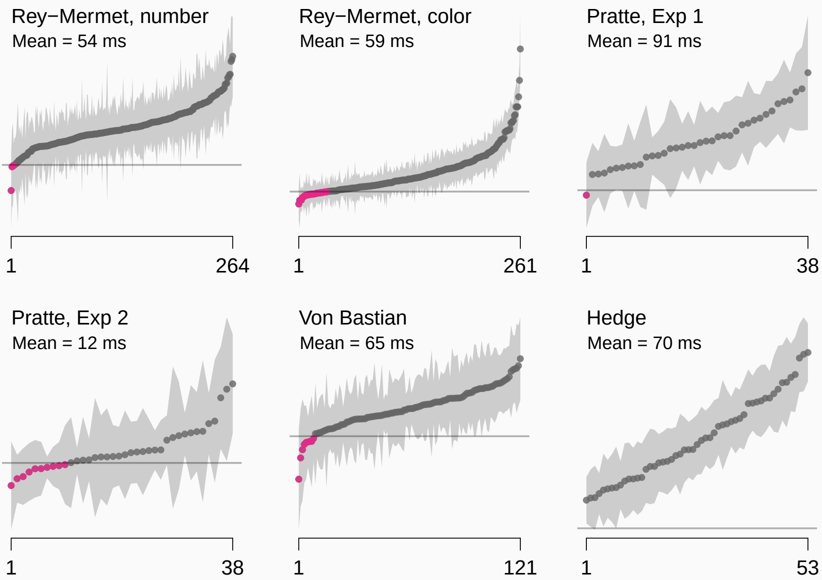


The Research Question

Does everyone show a true effect in the same direction?

Is this an interesting question?

Is this an interesting question?



Let's Do Some Modeling

Let Y_{ijk} be the response time

- of the ith person, $i = 1, \ldots, I$,

Let's Do Some Modeling

Let Y_{ijk} be the response time

- of the ith person, $i = 1, \ldots, I$,
- in the jth condition, $j =$ congruent, incongruent,

Let's Do Some Modeling

Let Y_{ijk} be the response time

- of the ith person, $i = 1, \ldots, I$,
- in the jth condition, $j =$ congruent, incongruent,
- for the kth trial, $k = 1, \ldots, K_{ij}$.

Let’s Do Some Modeling

Let Y_{ijk} be the response time

- of the ith person, $i = 1, \dots, I$,
- in the jth condition, $j =$ congruent, incongruent,
- for the kth trial, $k = 1, \dots, K_{ij}$.

Then

$$Y_{ijk} \sim \text{Normal}(\mu + \alpha_i + x_j\theta_i, \sigma^2).$$

Let's Do Some Modeling

Let Y_{ijk} be the response time

- of the ith person, $i = 1, \ldots, I$,
- in the jth condition, $j =$ congruent, incongruent,
- for the kth trial, $k = 1, \ldots, K_{ij}$.

Then

$$Y_{ijk} \sim \text{Normal}(\mu + \alpha_i + x_j\theta_i, \sigma^2).$$

We care about the collection of θ_i.

Models on Individuals' True Effects

Three models are needed:

Models on Individuals' True Effects

Three models are needed:

1. No one shows a true Stroop effect.

Models on Individuals' True Effects

Three models are needed:

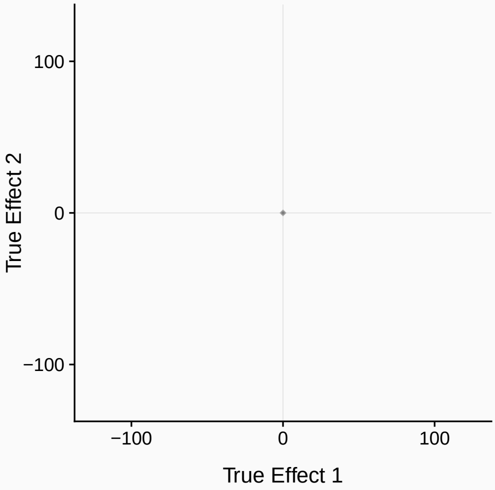
1. No one shows a true Stroop effect.
2. Everyone show a true Stroop effect in the same direction.

Models on Individuals' True Effects

Three models are needed:

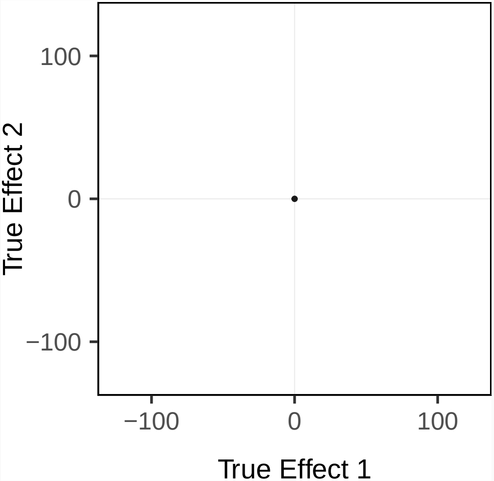
1. No one shows a true Stroop effect.
2. Everyone show a true Stroop effect in the same direction.
3. True Stroop effects vary in direction and size.

Models for Individual Differences and Ordinal Constraint



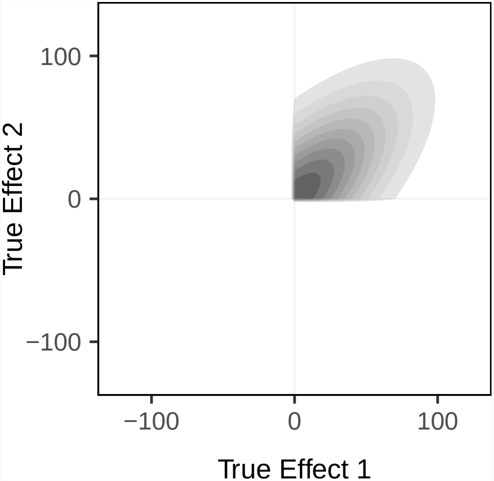
Haaf & Rouder (2017)

No One Shows a True Stroop Effect



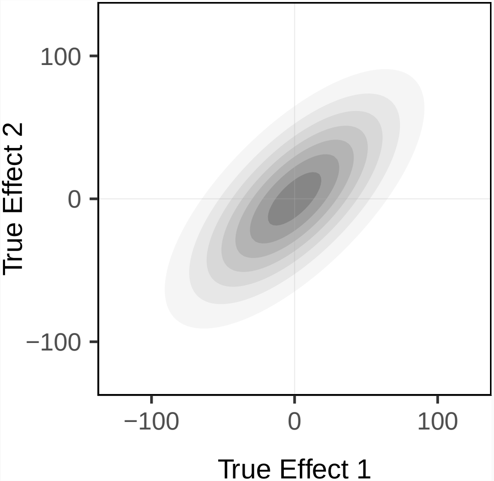
Haaf & Rouder (2017)

Everyone Show a True Stroop Effect in the Same Direction



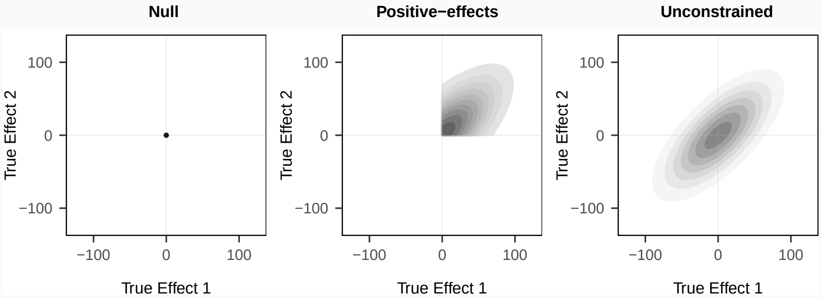
Haaf & Rouder (2017)

True Stroop Effects Vary in Direction and Size



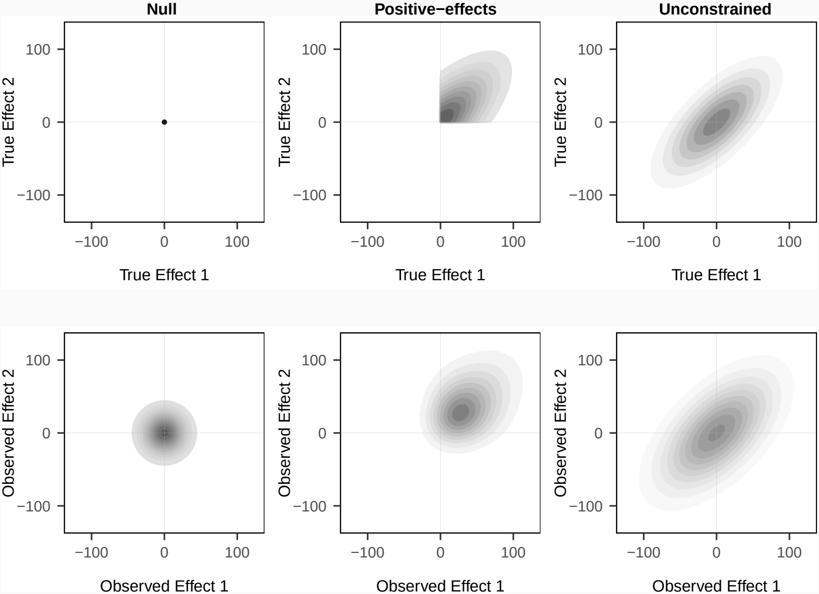
Haaf & Rouder (2017)

From Models. . .

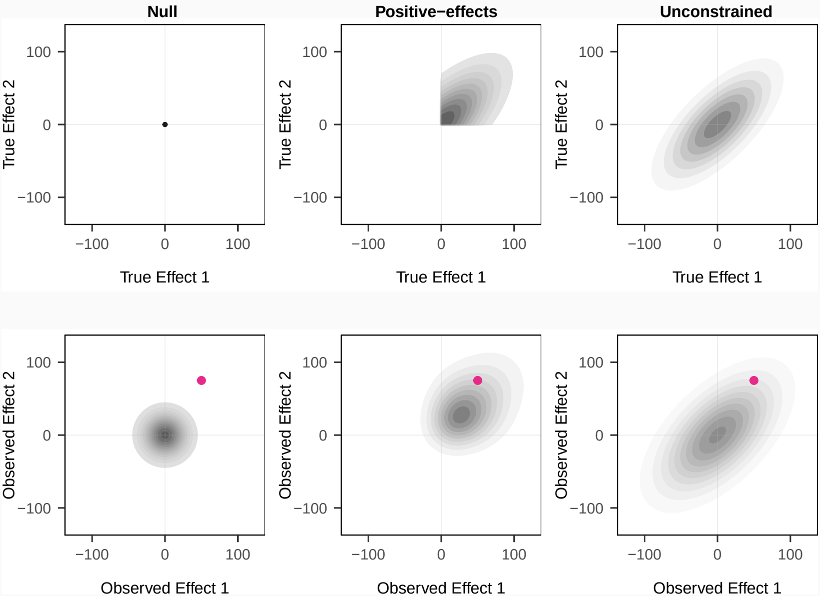


Null
Positive−effects
Unconstrained
100
0
−100
True Effect 2
True Effect 1

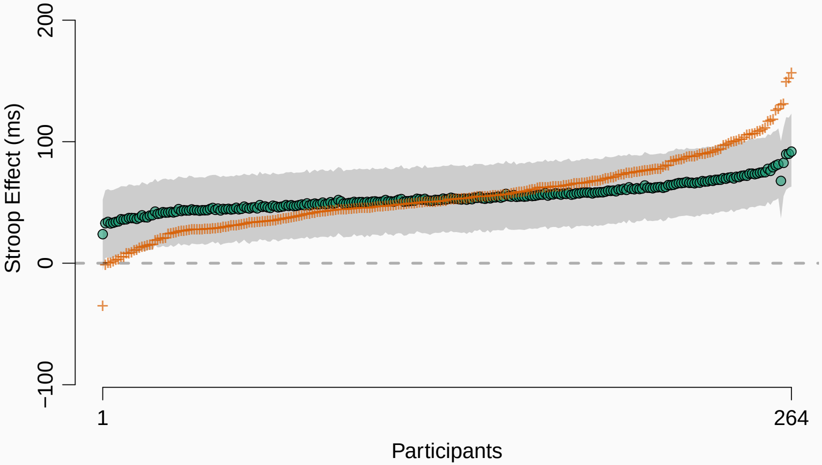
From Models... to Predictions



From Models... to Predictions... to Predictive Accuracy

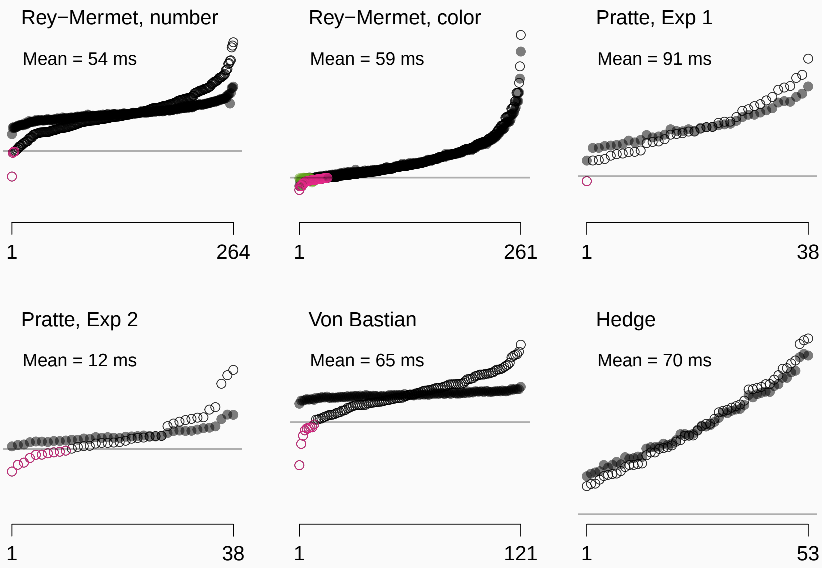


Ordinal Constraint for Individual Stroop Effects



Number-Stroop data by Rey-Mermet, Gade, & Oberauer (2018).

Does everyone Stroop?



Does everyone Stroop?

Study	Participants	Trials	Effect	BF_{pu}	BF_{p0}
Rey-Mermet, number	264	93	54 ms	**6.05**	$> 10^{300}$
Rey-Mermet, color	261	94	59 ms	10^{-4}	$> 10^{300}$
Pratte, Exp 1	38	146	90 ms	**4.75**	10^{73}
Pratte, Exp 2	38	165	12 ms	0.28	78.49
Von Bastian	121	46	64 ms	**6.38**	10^{62}
Hedge	53	410	69 ms	**6.33**	$> 10^{300}$

Does everyone Stroop?

Study	Participants	Trials	Effect	BF_{pu}	BF_{p0}
Rey-Mermet, number	264	93	54 ms	**6.05**	$> 10^{300}$
Rey-Mermet, color	261	94	59 ms	10^{-4}	$> 10^{300}$
Pratte, Exp 1	38	146	90 ms	**4.75**	10^{73}
Pratte, Exp 2	38	165	12 ms	0.28	78.49
Von Bastian	121	46	64 ms	**6.38**	10^{62}
Hedge	53	410	69 ms	**6.33**	$> 10^{300}$

Does everyone Stroop?

Study	Participants	Trials	Effect	BF_{pu}	BF_{p0}
Rey-Mermet, number	264	93	54 ms	**6.05**	$> 10^{300}$
Rey-Mermet, color	261	94	59 ms	10^{-4}	$> 10^{300}$
Pratte, Exp 1	38	146	90 ms	**4.75**	10^{73}
Pratte, Exp 2	**38**	**165**	**12 ms**	**0.28**	**78.49**
Von Bastian	121	46	64 ms	**6.38**	10^{62}
Hedge	53	410	69 ms	**6.33**	$> 10^{300}$

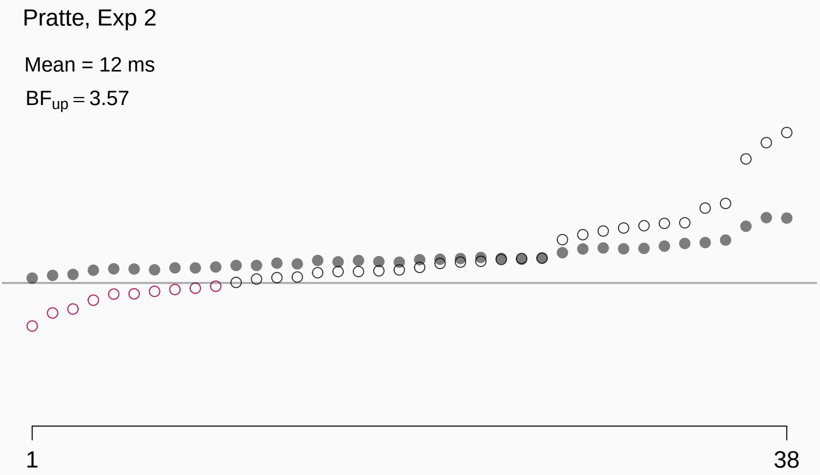
Motivated by Pratte et al., Experiment 2

- Why do some people not show a Stroop effect?

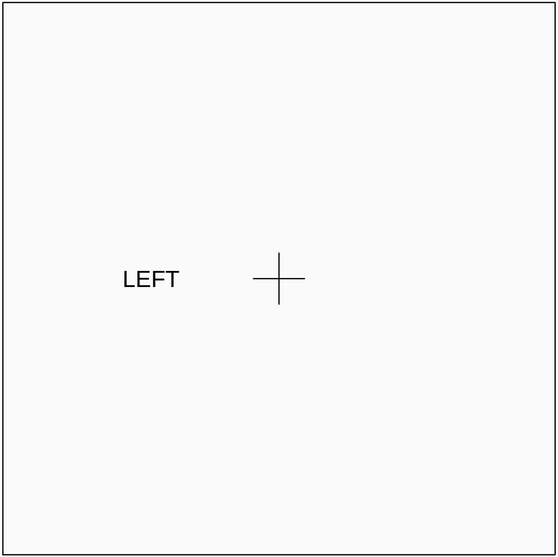
Motivated by Pratte et al., Experiment 2

- Why do some people not show a Stroop effect?
- What is the nature of their non-Stroopiness?

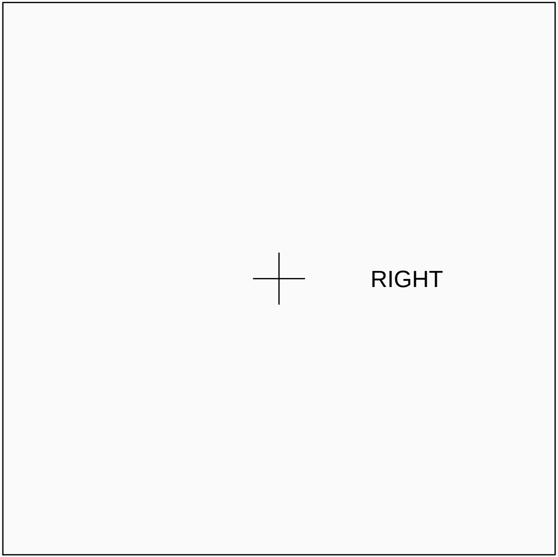
Motivated by Pratte et al., Experiment 2



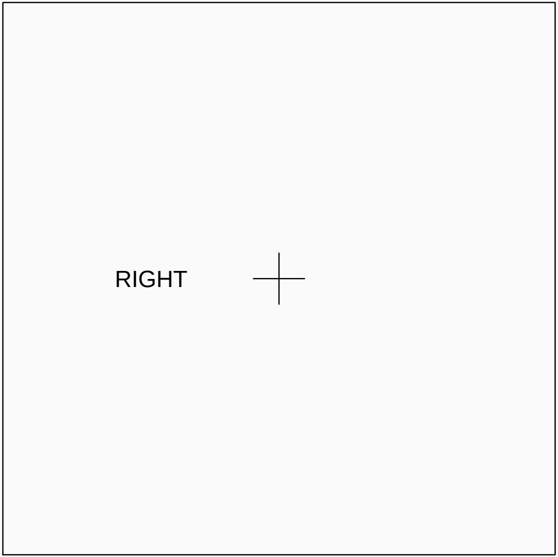
What could be different with the location Stroop?



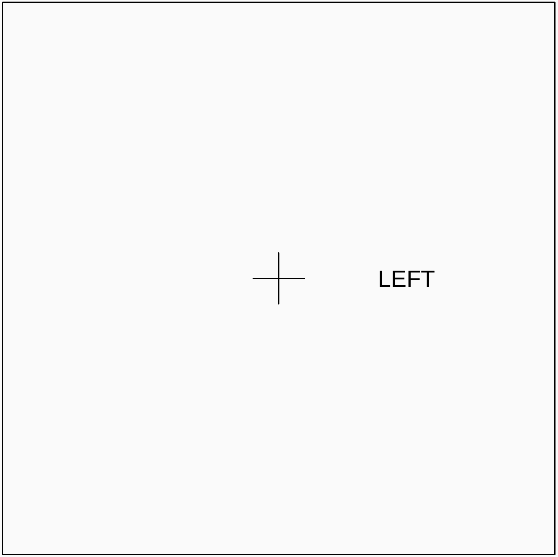
What could be different with the location Stroop?



What could be different with the location Stroop?



What could be different with the location Stroop?



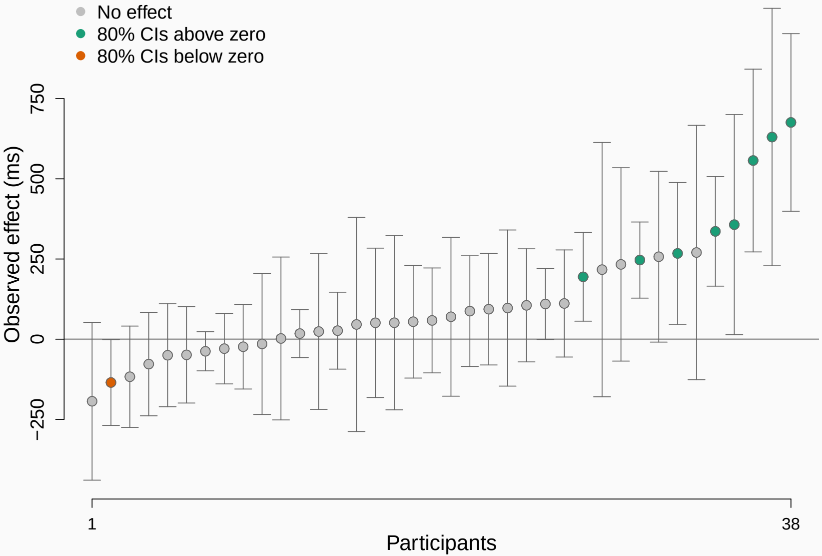
What could be different with the location Stroop?

- Different participants use different strategies.

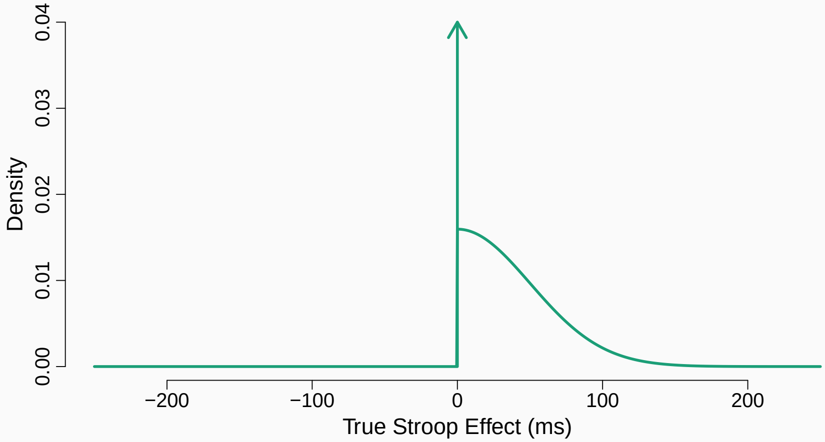
What could be different with the location Stroop?

- Different participants use different strategies.
- How would we tell?

How would we tell?

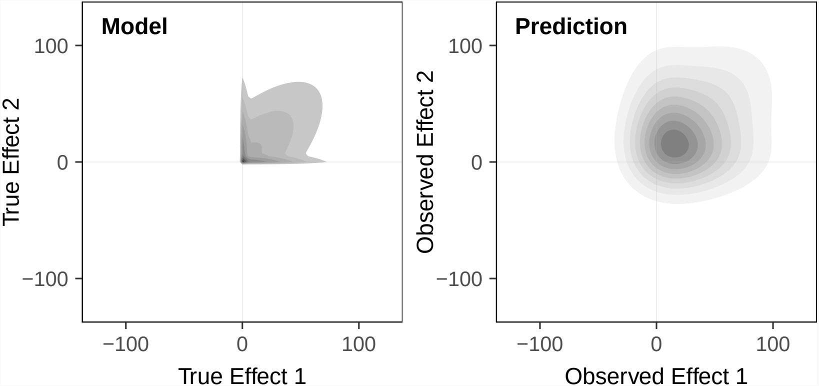


Some Do Some Don't



Haaf & Rouder (2019)

Modeling Approach: Some Do Some Don't Model



Haaf & Rouder (2019)

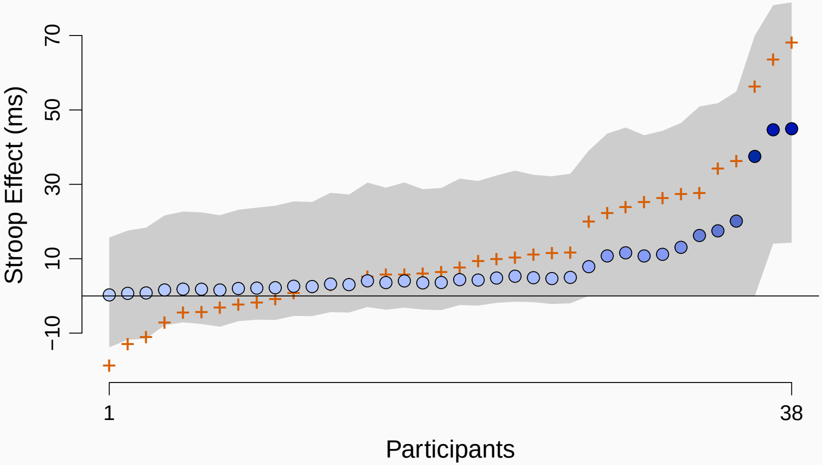
Location Stroop Results

- The Bayes factor of the some-do-some-don't model over the unconstrained model is 4.4 to 1.

Location Stroop Results

- The Bayes factor of the some-do-some-don't model over the unconstrained model is 4.4 to 1.
- Which participants do and which don't?

Classification Based on the Hierarchical Model



Haaf & Rouder (2019)

Summary

- Quantitative vs. qualitative individual differences.

Summary

- Quantitative vs. qualitative individual differences.
- If everyone shows an effect in the same direction we can rely on simple, universal theoretical accounts.

Summary

- Quantitative vs. qualitative individual differences.
- If everyone shows an effect in the same direction we can rely on simple, universal theoretical accounts.
- If some individuals robustly show no effect it can have implications for experimental methods and theory.

Summary

- Quantitative vs. qualitative individual differences.
- If everyone shows an effect in the same direction we can rely on simple, universal theoretical accounts.
- If some individuals robustly show no effect it can have implications for experimental methods and theory.
- If some individuals robustly show an opposite effect we may prefer more complex theories.

Summary

- Quantitative vs. qualitative individual differences.
- If everyone shows an effect in the same direction we can rely on simple, universal theoretical accounts.
- If some individuals robustly show no effect it can have implications for experimental methods and theory.
- If some individuals robustly show an opposite effect we may prefer more complex theories.
- In this case, answering *why* there are qualitative individual differences is key for theory development.

Do it yourself: The Does-Everyone t-Test

Function `quid()`: The Qualitative-Individual-Differences Test Function

- `quid()` is available on github:
  https://github.com/jstbcs/play/

Function `quid()`: The Qualitative-Individual-Differences Test Function

- `quid()` is available on github: https://github.com/jstbcs/play/
- The function may be installed as follows:

```
install.packages(c("BayesFactor", "MCMCpack", "curl"))
filename <- curl::curl("https://bit.ly/2ZqGOik")
source(filename)
```

Function `quid()`: The Qualitative-Individual-Differences Test Function

- Let's look at some data: Von Bastian, Souza, & Gade (2015), Experiment 1.

Function `quid()`: The Qualitative-Individual-Differences Test Function

- Let's look at some data: Von Bastian, Souza, & Gade (2015), Experiment 1.
- For the code to run you need the package `devtools` (`install.packages("devtools")`).

Function `quid()`: The Qualitative-Individual-Differences Test Function

- Let's look at some data: Von Bastian et al. (2015), Experiment 1.
- For the code to run you need the package `devtools` (`install.packages("devtools")`).

```
devtools::source_url("https://bit.ly/2SxyKtq")
```

Function `quid()`: The Qualitative-Individual-Differences Test Function

- Let's look at some data: Von Bastian et al. (2015).
- For the code to run you need the package `devtools` (`install.packages("devtools")`).

```
devtools::source_url("https://bit.ly/2SxyKtq")
```

The data are now loaded as a data frame called `stroop`.

Function `quid()`: The Qualitative-Individual-Differences Test Function

- The participant identification number is the variable `stroop$ID` (from 1 to 121).

Function `quid()`: The Qualitative-Individual-Differences Test Function

- The participant identification number is the variable `stroop$ID` (from 1 to 121).
- The condition is the variable `stroop$cond` (values are 1 for congruent and 2 for incongruent)

Function `quid()`: The Qualitative-Individual-Differences Test Function

- The participant identification number is the variable `stroop$ID` (from 1 to 121).
- The condition is the variable `stroop$cond` (values are 1 for congruent and 2 for incongruent)
- The response time for each trial in seconds is stored in the variable `stroop$rt`

Function `quid()`: The Qualitative-Individual-Differences Test Function

Now we are ready to use `quid()`:

```
res <- quid(id = stroop$ID
            , condition = stroop$cond
            , rt = stroop$rt)
```

Function `quid()`: The Qualitative-Individual-Differences Test Function

The output, `res`, is a list containing

- The posterior-mean estimates from the unconstrained model for each individuals' effect (θ_i).

Function `quid()`: The Qualitative-Individual-Differences Test Function

The output, `res`, is a list containing

- The posterior-mean estimates from the unconstrained model for each individuals' effect (θ_i).
- The posterior overall effect.

Function `quid()`: The Qualitative-Individual-Differences Test Function

The output, `res`, is a list containing

- The posterior-mean estimates from the unconstrained model for each individuals' effect (θ_i).
- The posterior overall effect.
- The posterior standard deviation of estimated effects.

Function `quid()`: The Qualitative-Individual-Differences Test Function

The output, `res`, is a list containing

- The posterior-mean estimates from the unconstrained model for each individuals' effect (θ_i).
- The posterior overall effect.
- The posterior standard deviation of estimated effects.
- The Bayes factors for the models, and raw outputs from the underlying MCMC chains (`res$bfs`).

Function `quid()`: The Qualitative-Individual-Differences Test Function

The output, `res`, is a list containing

- The posterior-mean estimates from the unconstrained model for each individuals' effect (θ_i).
- The posterior overall effect.
- The posterior standard deviation of estimated effects.
- The Bayes factors for the models, and raw outputs from the underlying MCMC chains (`res$bfs`).

```
res$bfs
```

```
##       bf.1u        bf.pu        bf.0u
## 9.777135e-01 6.310874e+00 1.002935e-62
```

Function `quid()`: Prior Settings

- An important input to `quid()` is `prior`.

Function `quid()`: Prior Settings

- An important input to `quid()` is `prior`.
- The default is `prior = c(1/6, 1/10)`.

Function `quid()`: Prior Settings

- An important input to `quid()` is `prior`.
- The default is `prior = c(1/6, 1/10)`.
- These values are scales on:

Function `quid()`: Prior Settings

- An important input to `quid()` is `prior`.
- The default is `prior = c(1/6, 1/10)`.
- These values are scales on:

(a) about how large we expect the mean effect to be, and

Function `quid()`: Prior Settings

- An important input to `quid()` is `prior`.
- The default is `prior = c(1/6, 1/10)`.
- These values are scales on:

(a) about how large we expect the mean effect to be, and

(b) about how much we expect individuals to differ from this mean effect.

Function `quid()`: Prior Settings

- An important input to `quid()` is `prior`.
- The default is `prior = c(1/6, 1/10)`.

```
largeVals <- c(80/200, 40/200)
resB <- quid(id = stroop$ID
             , condition = stroop$cond
             , rt = stroop$rt
             , prior = largeVals)
resB$bfs
```

```
##        bf.1u        bf.pu        bf.0u
## 3.077463e+00 4.664942e+00 2.570078e-62
```

Function `quid()`: Other Priors

$$(\mu, \sigma^2) \propto \frac{1}{\sigma^2},$$

$$\alpha_i \sim \text{Normal}(0, g_\alpha \sigma^2),$$

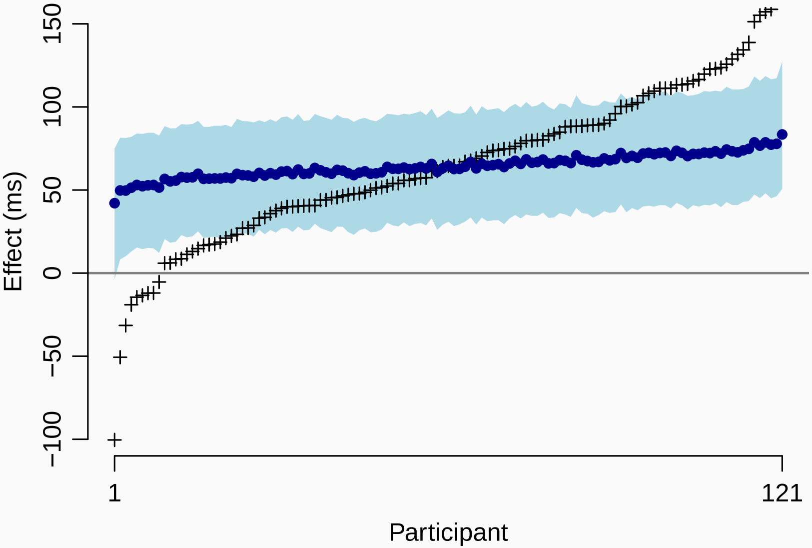
Most important priors:

$$\theta_i \sim \text{Normal}(\mu_\theta, g_\theta \sigma^2),$$

$$\mu_\theta \sim \text{Normal}(0, g_{\mu\theta}, \sigma^2).$$

gs have scaled χ^2-distributions, and the scales are set by `prior`.

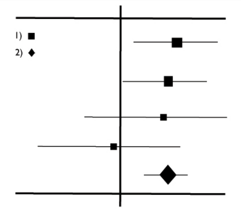
Function `quid()`: You can make nice plots!



Does every study show an effect in the expected direction?

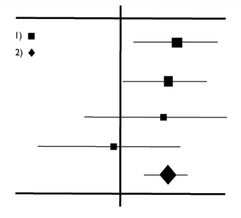
Why This Is A Good Question For Meta-analysis

- The usual meta-analytic question: What is the overall effect combined over a bunch of studies?



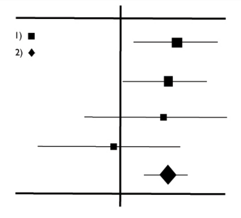
Why This Is A Good Question For Meta-analysis

- The usual meta-analytic question: What is the overall effect combined over a bunch of studies?
- The overall effect depends on many things.



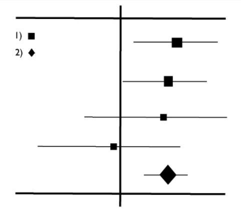
Why This Is A Good Question For Meta-analysis

- The usual meta-analytic question: What is the overall effect combined over a bunch of studies?
- The overall effect depends on many things.
- Choices of paradigms and variables.



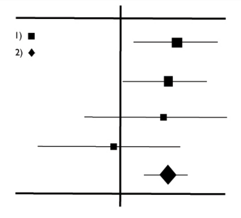
Why This Is A Good Question For Meta-analysis

- The usual meta-analytic question: What is the overall effect combined over a bunch of studies?
- The overall effect depends on many things.
- Choices of paradigms and variables.
- What is currently hot in the field.



Why This Is A Good Question For Meta-analysis

- The usual meta-analytic question: What is the overall effect combined over a bunch of studies?
- The overall effect depends on many things.
- Choices of paradigms and variables.
- What is currently hot in the field.
- If the target contrast is robust the *direction* of the effect should not be affected.

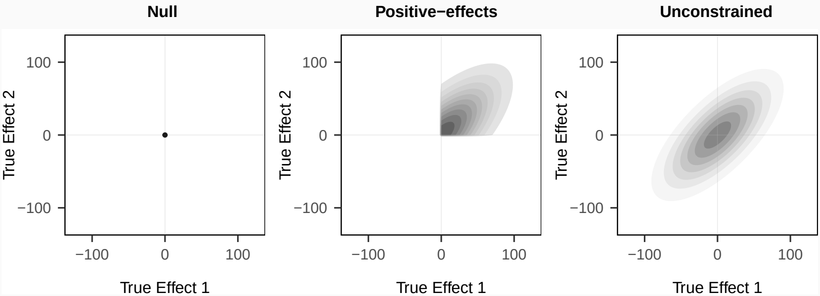


New Meta-Analytic Question

Does Every Study In A Collection Plausibly Show an Effect in the Same Direction?

(Haaf, 2018; Rouder, Haaf, Davis-Stober, & Hilgard, 2019)

Meta-Analytic Models



Application: Early language development

- Do toddlers recognize familiar words?

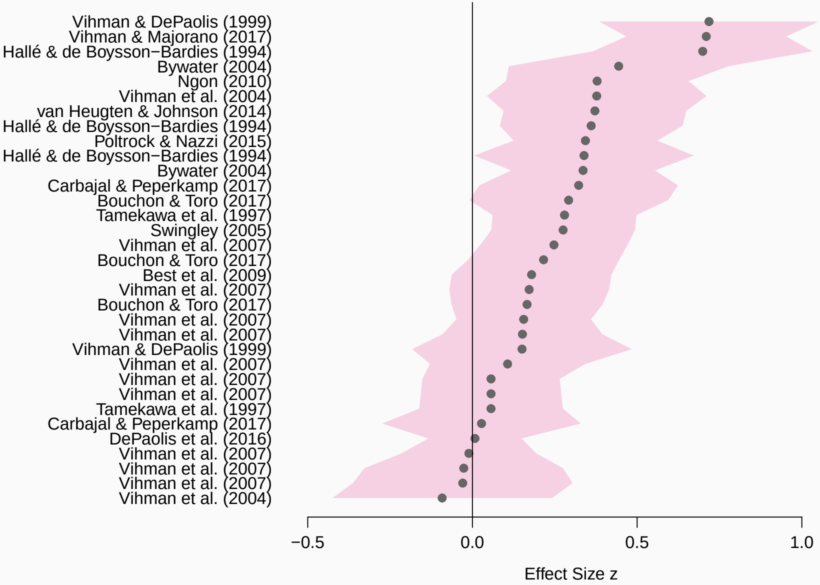
Application: Early language development

- Do toddlers recognize familiar words?
- General finding: Toddlers (~11-20 months) pay longer attention to familiar words than novel ones.

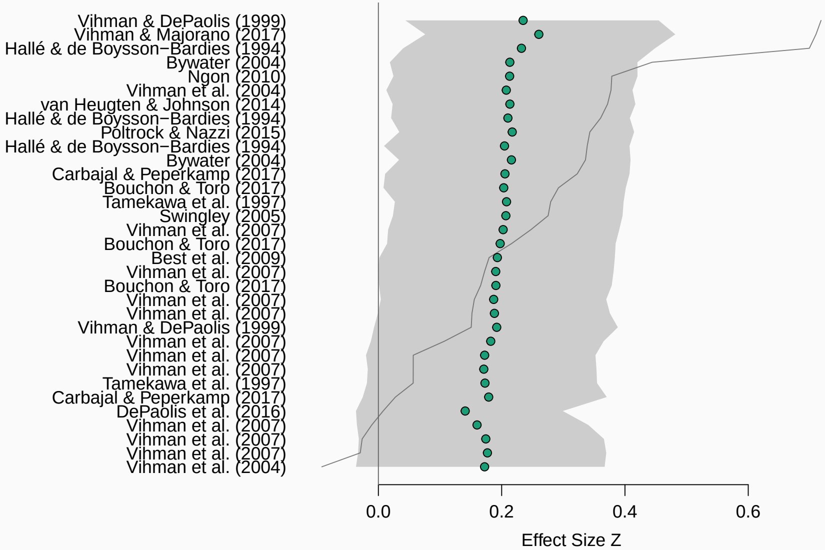
Application: Early language development

- Do toddlers recognize familiar words?
- General finding: Toddlers (∼11-20 months) pay longer attention to familiar words than novel ones.
- Carbajal (2018) conducted a meta-analysis with 33 studies.

Application: Early language development



Application: Early language development



Application: Early language development

- The Bayes factor of the every-study-does over the unconstrained model is 8.01 to 1.

Application: Early language development

- The Bayes factor of the every-study-does over the unconstrained model is 8.01 to 1.
- The Bayes factor of the every-study-does over the null model is 4.83 to 1.

Summary II

- Evidence that every study shows the familiar-words effect.

Summary II

- Evidence that every study shows the familiar-words effect.
- The average effect size is 0.2 (Fisher's Z).

Summary II

- Evidence that every study shows the familiar-words effect.
- The average effect size is 0.2 (Fisher's Z).
- Qualitative interactions (Gail & Simon, 1985).

Summary II

- Evidence that every study shows the familiar-words effect.
- The average effect size is 0.2 (Fisher's Z).
- Qualitative interactions (Gail & Simon, 1985).
- Does-every-study approach is now implemented in the `metaBMA` package in `R`.

Back to Individual Differences

- Cognitive Psychology is more complicated than the Stroop effect.

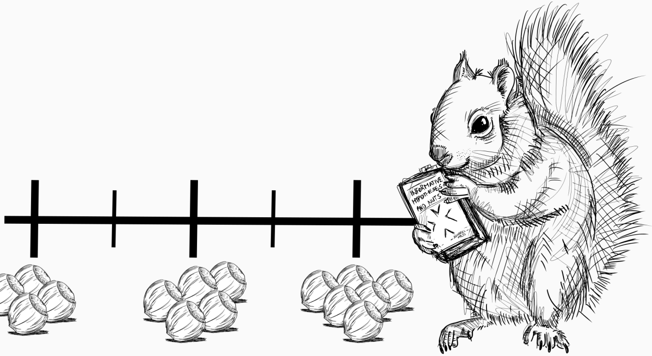
Back to Individual Differences

- Cognitive Psychology is more complicated than the Stroop effect.
- Developing individual differences approaches for more diverse data patterns.

Example: How do we represent numbers internally?

How do we represent numbers internally?

1. Analog representation.



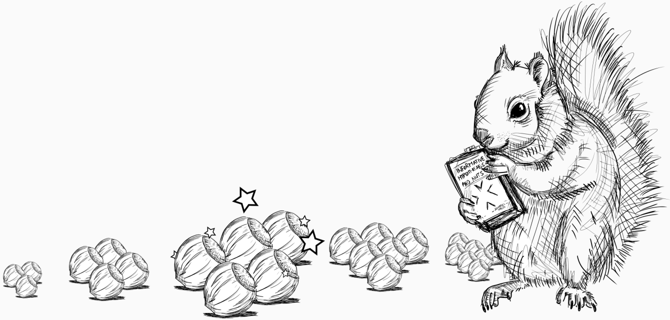
How do we represent numbers internally?

1. Analog representation.
2. Propositional representation.

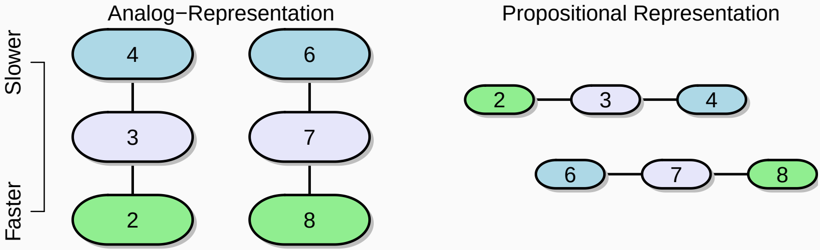
 <  = true

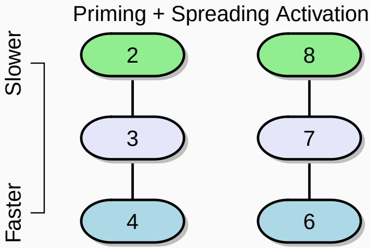
How do we represent numbers internally?

1. Analog representation.
2. Propositional representation.
3. Priming + spreading activation.

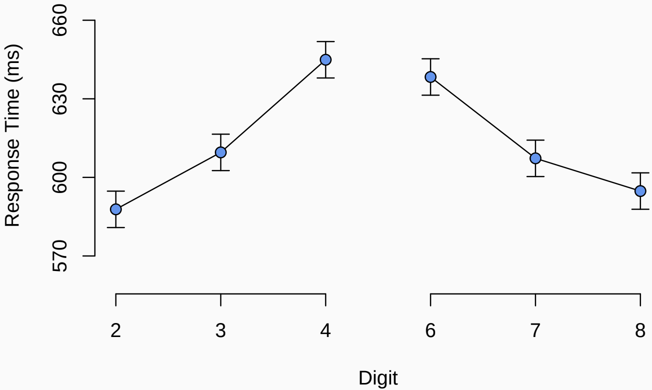


Theoretical positions as ordinal models





Rouder, Lu, Speckman, Sun, & Jiang (2005)



Individual Differences in Number Representation

Does everyone represent numbers the same way?

Individual Differences in Number Representation

Does everyone represent numbers the same way?

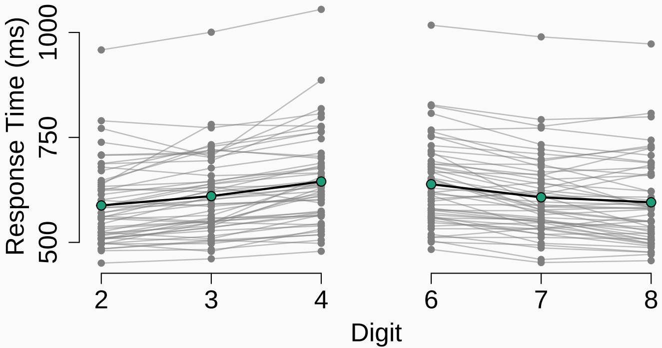
- Common mechanism $\rightarrow$ common processing architecture.

Individual Differences in Number Representation

Does everyone represent numbers the same way?

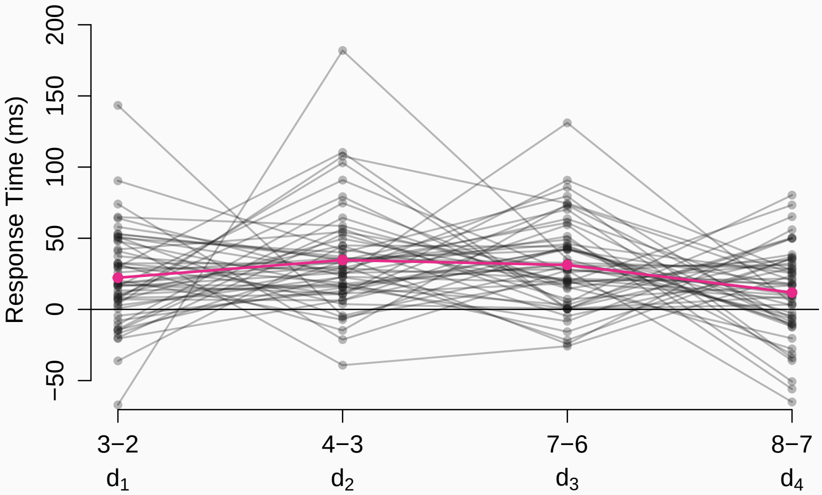
- Common mechanism $\rightarrow$ common processing architecture.
- Mixture of representations $\rightarrow$ what is the underlying mechanism?

Individual Differences in Number Representation

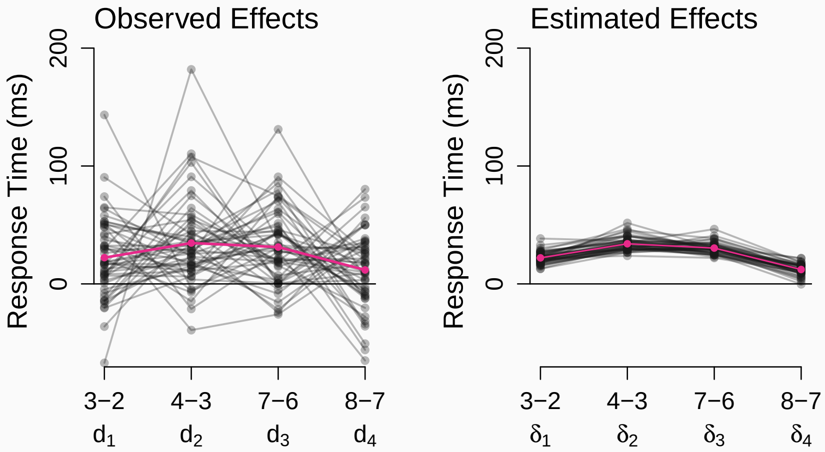


Data by Rouder, Lu, Speckman, Sun, & Jiang (2005).

Individual Differences in Number Representation

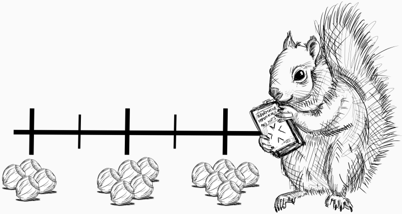


Individual Differences in Number Representation



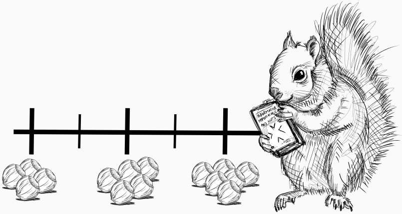
Individual Differences in Number Representation

- Preferred model: **Analog representation** model



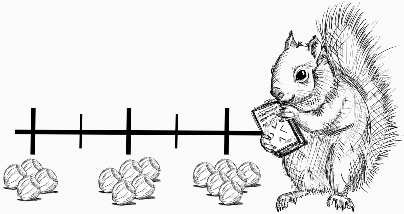
Individual Differences in Number Representation

- Preferred model: **Analog representation** model
- Preferred 9.78-to-1 over the **None of the above** model



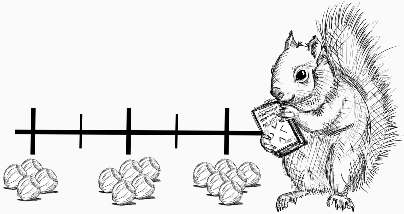
Individual Differences in Number Representation

- Preferred model: **Analog representation** model
- Preferred 9.78-to-1 over the **None of the above** model
- Preferred 3×10^{55}-to-1 over the **Propositional representation** model



Individual Differences in Number Representation

- Preferred model: **Analog representation** model
- Preferred 9.78-to-1 over the **None of the above** model
- Preferred 3×10^{55}-to-1 over the **Propositional representation** model
- Bayes factor for **Priming + spreading activation** model cannot be estimated



Take Away

Does everyone show an effect in the same direction?

- To answer this question we need a combination of new methods and theoretical considerations.

Take Away

Does everyone show an effect in the same direction?

- To answer this question we need a combination of new methods and theoretical considerations.
- Everyone Stroops.

Take Away

Does everyone show an effect in the same direction?

- To answer this question we need a combination of new methods and theoretical considerations.
- Everyone Stroops.
- Qualitative vs. quantitative individual differences is a useful distrinction in cognitive psychology.

Take Away

Does everyone show an effect in the same direction?

- To answer this question we need a combination of new methods and theoretical considerations.
- Everyone Stroops.
- Qualitative vs. quantitative individual differences is a useful distrinction in cognitive psychology.
- For individual differences research: Assessing how individuals vary without overstating individual differences.

Take Away

Does everyone show an effect in the same direction?

- To answer this question we need a combination of new methods and theoretical considerations.
- Everyone Stroops.
- Qualitative vs. quantitative individual differences is a useful distrinction in cognitive psychology.
- For individual differences research: Assessing how individuals vary without overstating individual differences.
- We first need to know *that people have a similar processing architecture* before we can report average effects.

Thank you!

Are there any questions?

Normal Models?!?

- But RT is skewed rather than symmetric.

Normal Models?!?

- But RT is skewed rather than symmetric.

1. We care about effects on RT, and the models are relatively robust to violations on the trial-RT level.

Normal Models?!?

- But RT is skewed rather than symmetric.

1. We care about effects on RT, and the models are relatively robust to violations on the trial-RT level.

- Advantage of the normal specification:

Normal Models?!?

- But RT is skewed rather than symmetric.

1. We care about effects on RT, and the models are relatively robust to violations on the trial-RT level.

- Advantage of the normal specification:

1. The normal is computationally convenient

Normal Models?!?

- But RT is skewed rather than symmetric.

1. We care about effects on RT, and the models are relatively robust to violations on the trial-RT level.

- Advantage of the normal specification:

1. The normal is computationally convenient
2. The effect is easily parameterized and the placement of constraint is straightforward to implement.

References

Carbajal, M. J. (2018). *Separation and acquisition of two languages in earlychildhood: A multidisciplinary approach* (PhD thesis). Université Paris Sciences et Lettres.

Gail, M., & Simon, R. (1985). Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*, *41*(2), 361–372.

Haaf, J. M. (2018). *A hierarchical Bayesian analysis of multiple order constraints in behavioral science* (PhD thesis). University of Missouri.

Haaf, J. M., & Rouder, J. N. (2017). Developing constraint in Bayesian mixed models. *Psychological Methods*, *22*(4), 779–798.

Haaf, J. M., & Rouder, J. N. (2019). Some do and some don't? Accounting for variability of individual difference structures. *Psychonomic Bulletin and Review*, *26*, 772–789. Retrieved from https://doi.org/10.3758/s13423-018-1522-x

Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception & Psychophysics*, *72*, 2013–2025.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Retrieved from http://dx.doi.org/10.1037/xlm0000450

Rouder, J. N., Haaf, J. M., Davis-Stober, C. P., & Hilgard, J. (2019). Beyond overall effects: A Bayesian approach to finding constraints in meta-analysis. *Psychological Methods*.

Rouder, J. N., Lu, J., Speckman, P. L., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin and Review*, *12*, 195–223.

Von Bastian, C. C., Souza, A. S., & Gade, M. (2015). No evidence for bilingual cognitive advantages: A test of