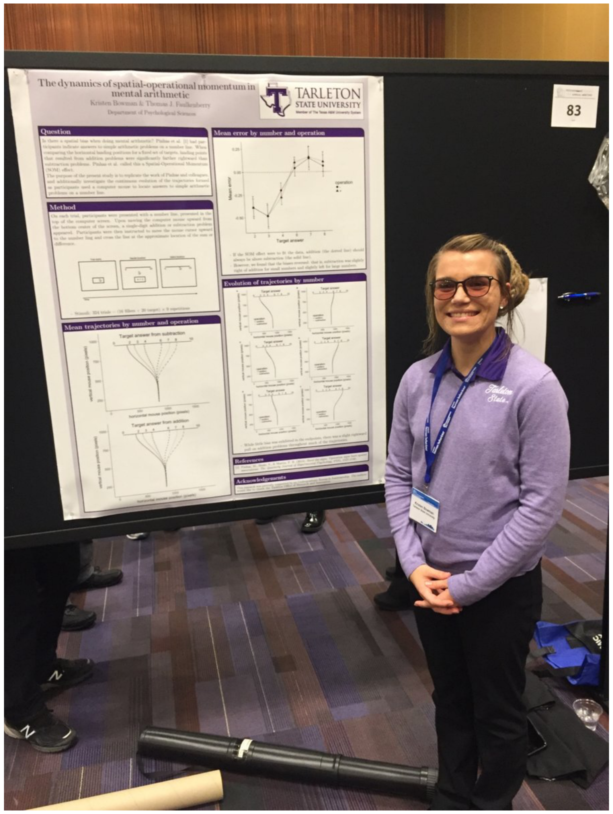
Modeling individual difference structures in the size congruity effect

Thomas J. Faulkenberry, Kristen A. Bowman, Sabrina Hetzel

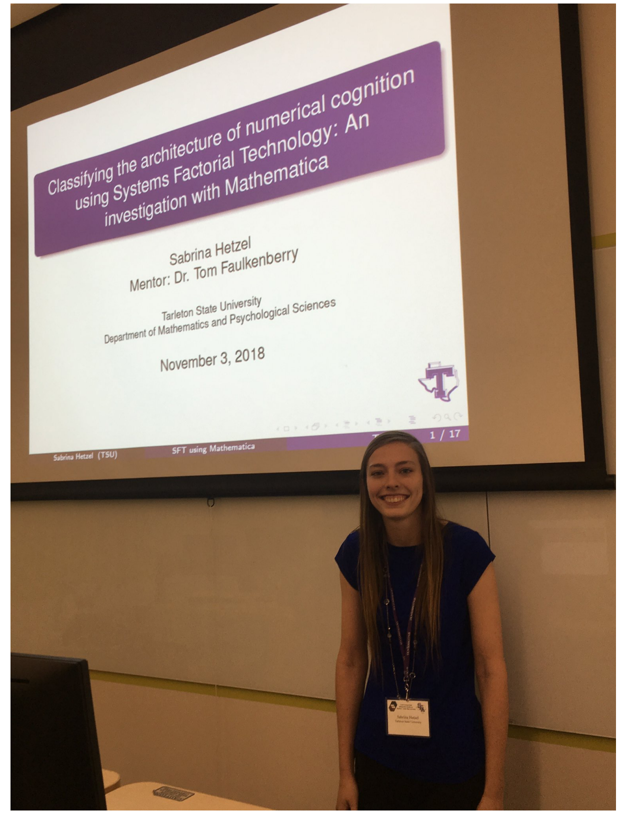
Tarleton State University

Psychonomics 2018

My students



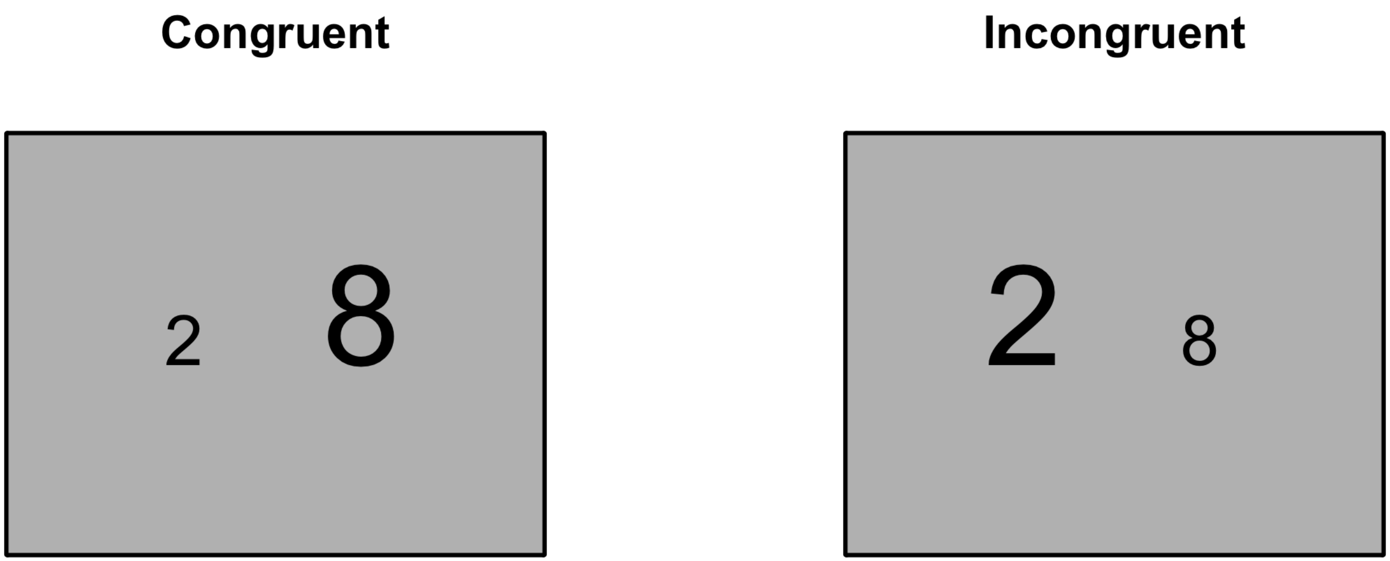
Kristen Bowman



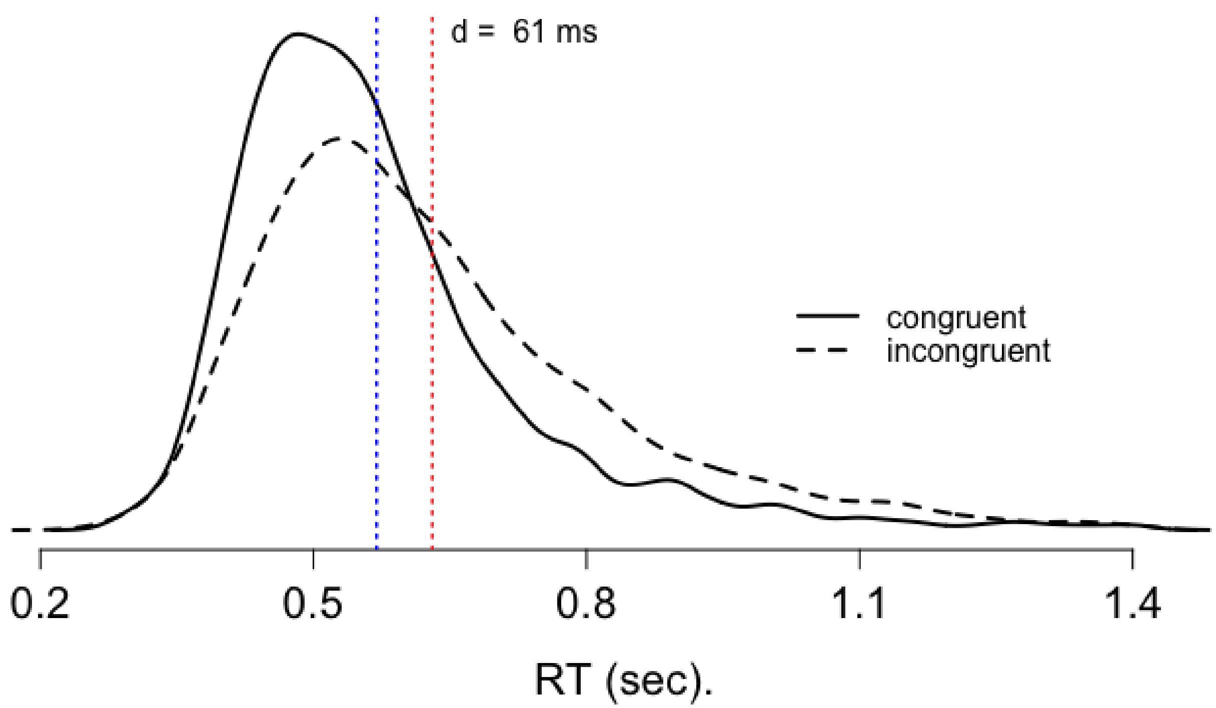
Sabrina Hetzel

Size congruity effect

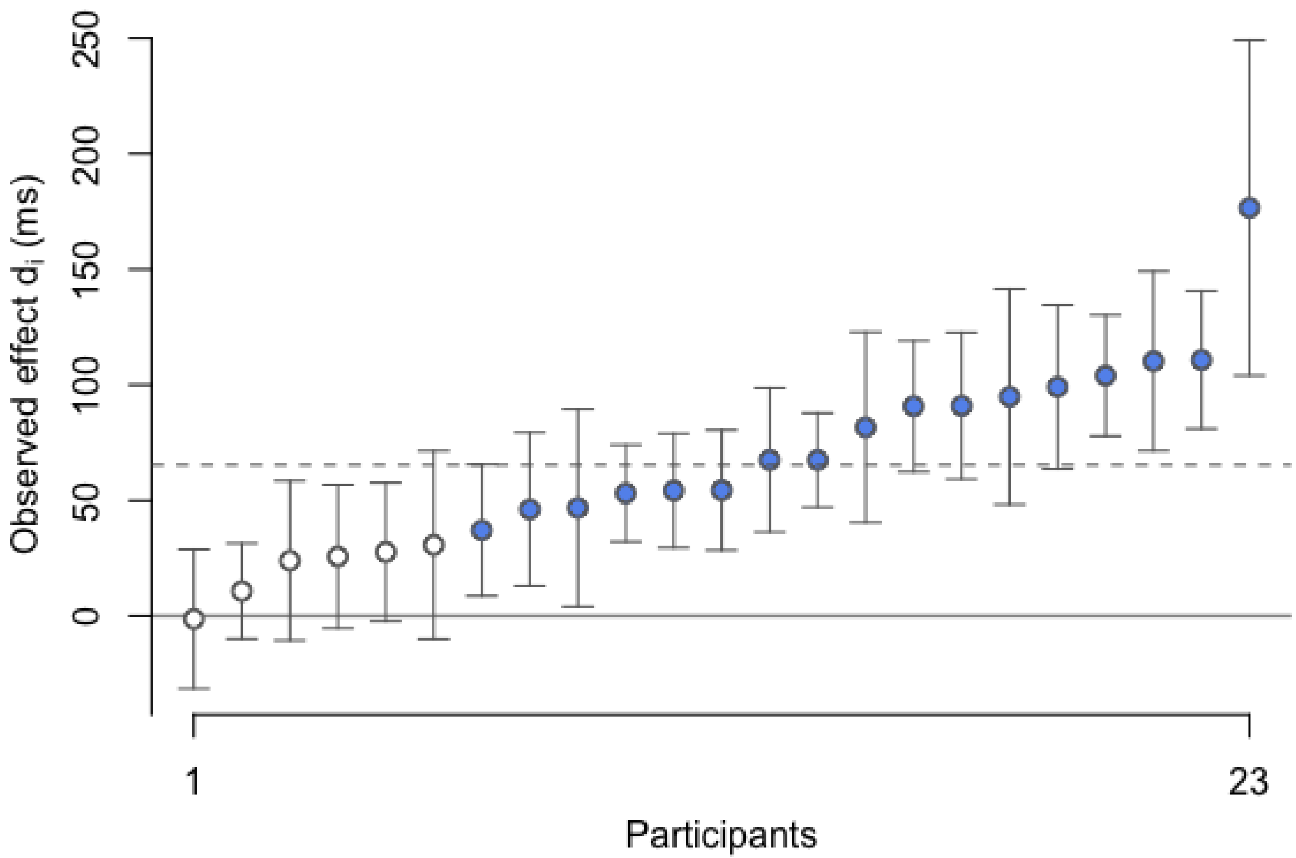
Typical laboratory task: choose the **physically larger** digit



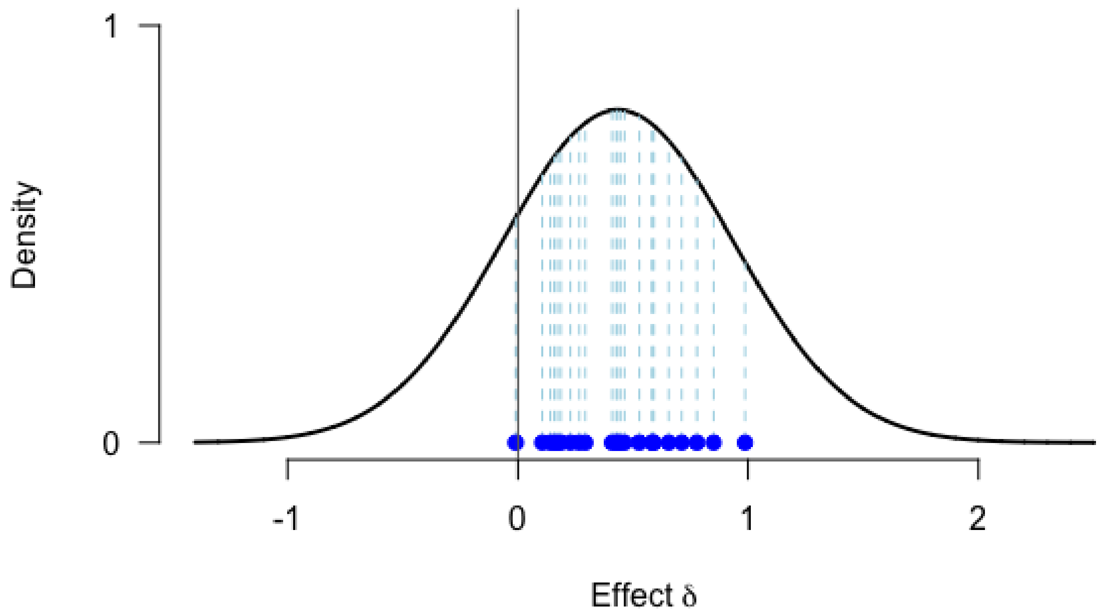
Typical result – mean RT larger for *incongruent trials*



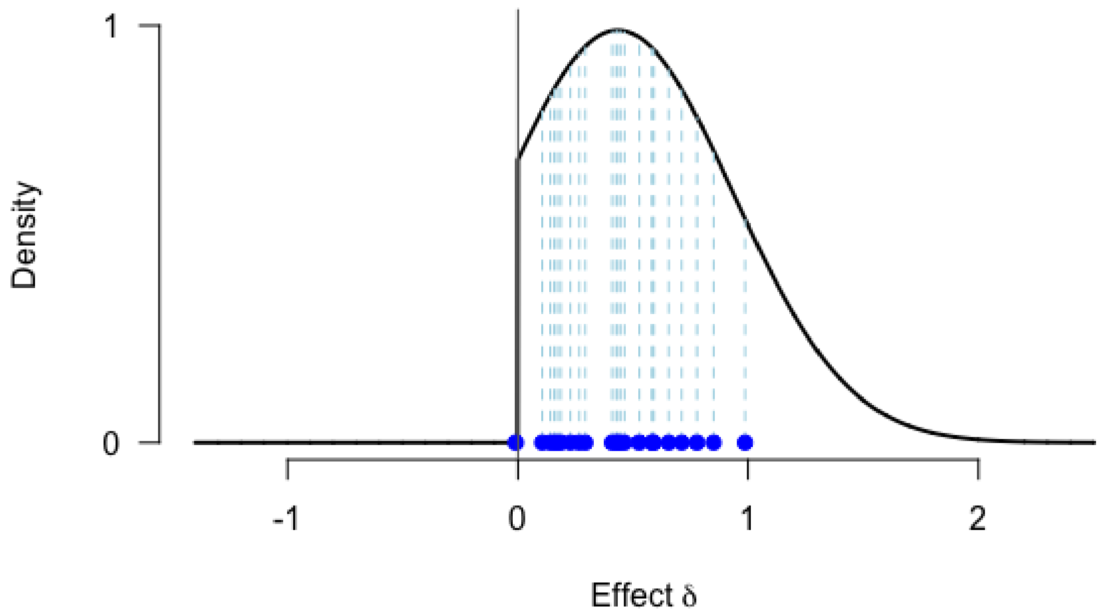
What about individual effects?



Suppose these observed effects d_i are drawn from population of *true* effects δ. What is the structure of this population?



Suppose these observed effects d_i are drawn from population of *true* effects δ. What is the structure of this population?



A new question

Does everybody *exhibit the size congruity effect?*

- if *yes*, then SCE is obligatory, resistant to strategic control, . . .
- if *no*, then SCE is complex, malleable, . . .

Importantly, both answers have downstream consequences for processing architecture of numerical cognition

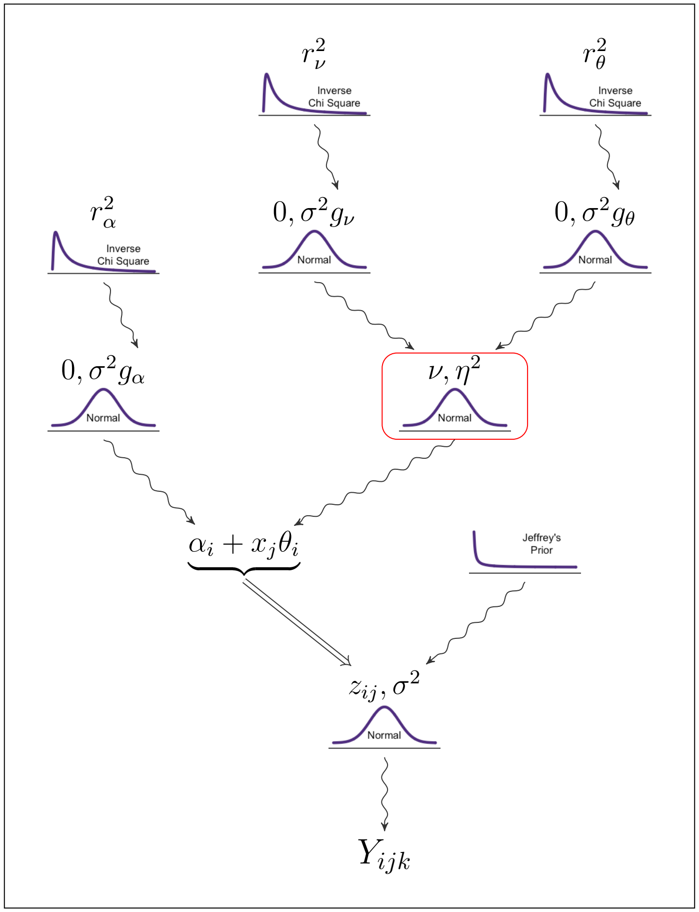
A new question

Does everybody *exhibit size congruity effect?*

How to answer:

- use the approach of Haaf and Rouder (2017) to build models of individual difference structures in SCE
- adjudicate the models via Bayesian model comparison

Hierarchical structure



Basic idea (Haaf & Rouder, 2017):

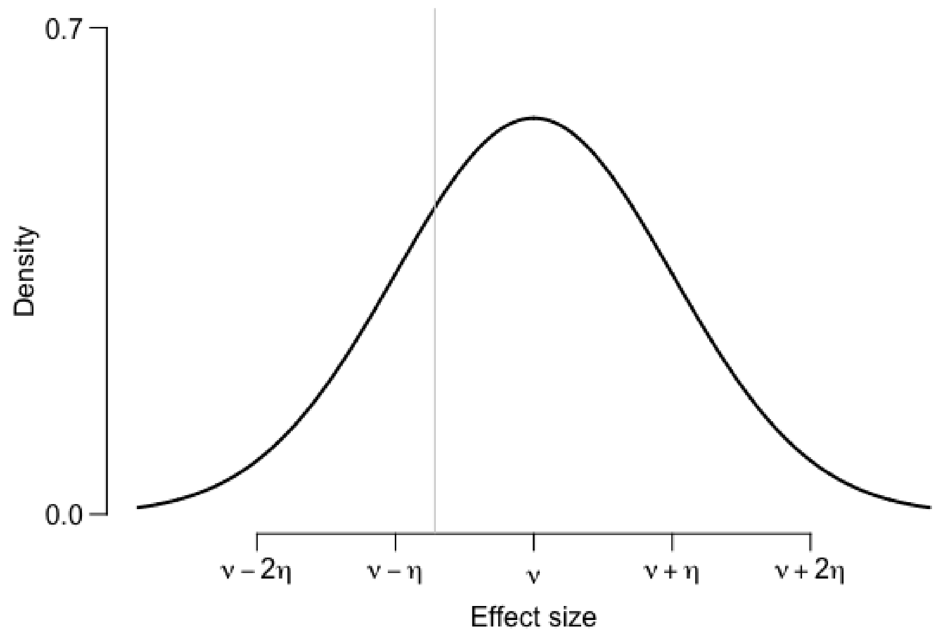
- model RTs as a random-effects linear model with effect parameter θ_i $(i = 1 \ldots, N)$
- assume (as baseline) that θ_i is drawn from a normal distribution with mean ν and variance η^2
- use g-priors (Zellner & Siow, 1980), specifying *a priori* scale on variance of overall effect and individual variability around effect
- define competing models by **constraining** effect parameter θ_i

Four competing models

1. Unrestricted model, $\mathcal{M}_u$
2. Positive-effects model, $\mathcal{M}_+$
3. Common-effect model, $\mathcal{M}_1$
4. Null-effect model, $\mathcal{M}_0$

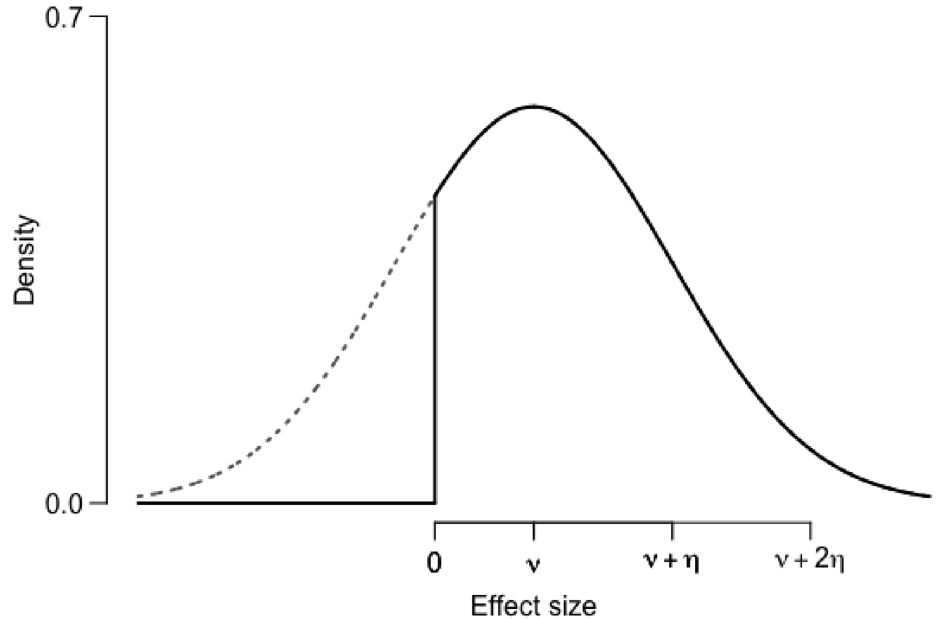
Unrestricted model

$\mathcal{M}_u : \theta_i \sim \text{Normal}(\nu, \eta^2)$



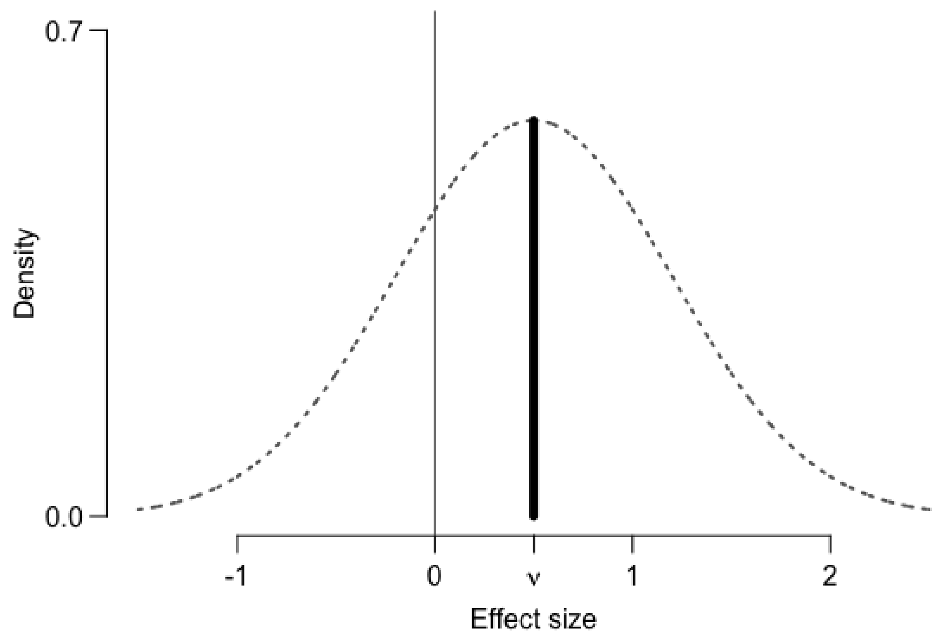
Positive-effects model

- $\mathcal{M}_+ : \theta_i \sim \text{Normal}_+(\nu, \eta^2)$



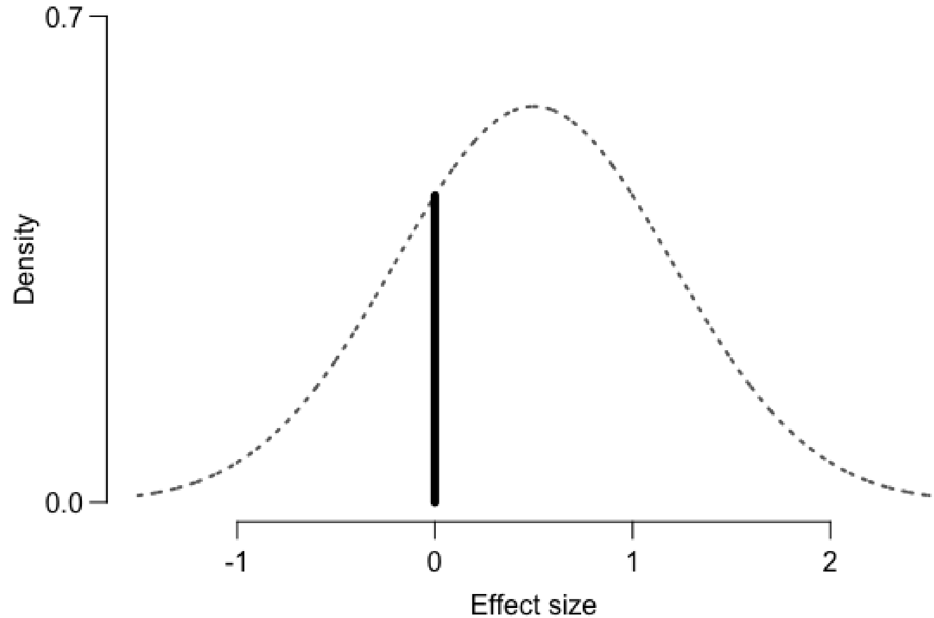
Common-effect model

- $\mathcal{M}_1 : \theta_i = \nu$



Null-effect model

- $\mathcal{M}_0 : \theta_i = 0$



Model comparison

We use the **Bayes factor**, which indexes how well the observed data $\boldsymbol{Y}$ are predicted under one model relative to another:

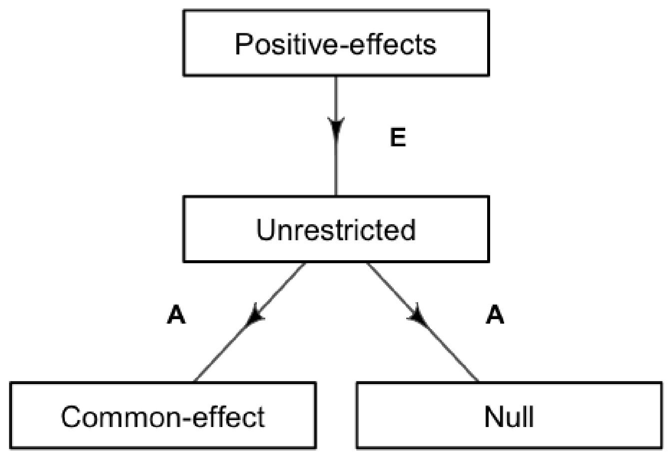
$$B_{ab} = \frac{p(\boldsymbol{Y} \mid \mathcal{M}_a)}{p(\boldsymbol{Y} \mid \mathcal{M}_b)}$$

Ex: suppose $B_{ab} = 10$. This means:

- the observed data are 10 times more likely under $\mathcal{M}_a$ compared to $\mathcal{M}_b$
- "10-to-1 evidence for $\mathcal{M}_a$ over $\mathcal{M}_b$"

Bayes factor computations

So $B_{ab} = \dfrac{p(\boldsymbol{Y} \mid \mathcal{M}_a)}{p(\boldsymbol{Y} \mid \mathcal{M}_b)}$. How do we compute that?



Bayes factor computations

So $B_{ab} = \dfrac{p(\boldsymbol{Y} \mid \mathcal{M}_a)}{p(\boldsymbol{Y} \mid \mathcal{M}_b)}$. How do we compute that?

$$p(\boldsymbol{Y} \mid \mathcal{M}) = \int_{\boldsymbol{\xi} \in \Xi} p(\boldsymbol{Y} \mid \boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi}$$

Bayes factor computations

So $B_{ab} = \dfrac{p(\boldsymbol{Y} \mid \mathcal{M}_a)}{p(\boldsymbol{Y} \mid \mathcal{M}_b)}$. How do we compute that?

$$p(\boldsymbol{Y} \mid \mathcal{M}) = \int_{\boldsymbol{\xi} \in \Xi} p(\boldsymbol{Y} \mid \boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi}$$

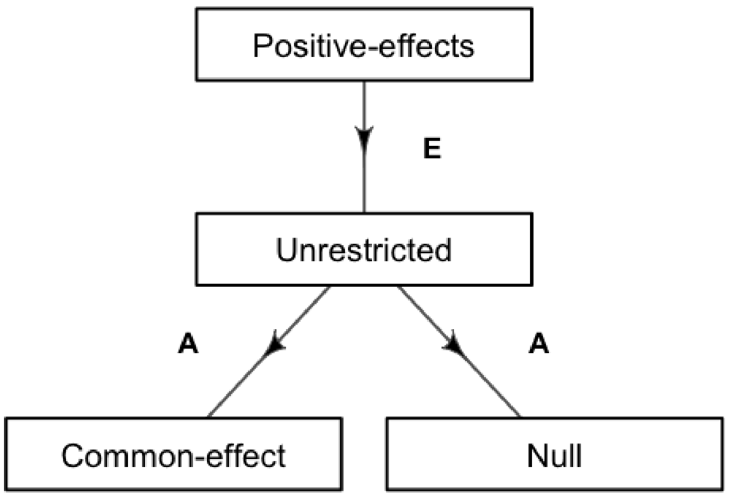
Problem: for our models $\mathcal{M}$, the parameter vectors $\boldsymbol{\xi}$ look like

$$\boldsymbol{\xi} = (\mu, \sigma^2, \nu, \alpha_1, \ldots, \alpha_N, \theta_1, \ldots, \theta_N, g_\alpha, g_\nu, g_\theta)$$

so the integral is carried out in $\mathbb{R}^{2N+6}$.

For $N = 35$, this would be a 76-dimensional integral!

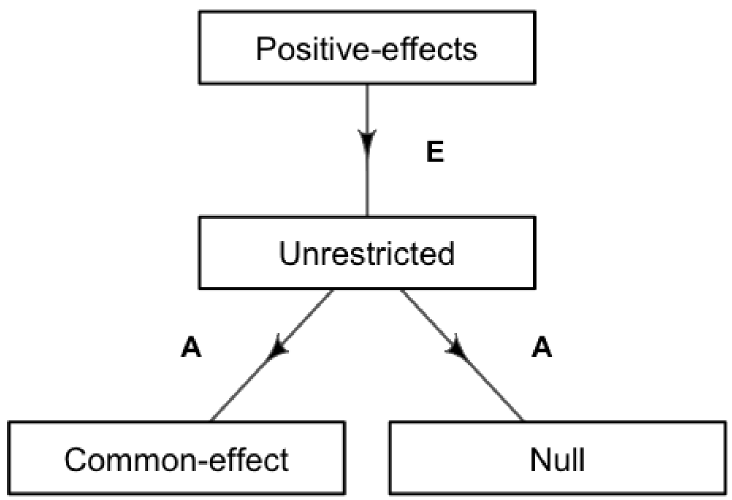
Bayes factor computations



A = analytic approach

- Zellner & Siow (1980); Rouder et al. (2012)
- place g-priors on individual intercepts and effect parameters
- everything *except* the g-parameters integrates symbolically
- g-parameters can be well approximated with MCMC sampling
- techniques coded into BayesFactor package in R

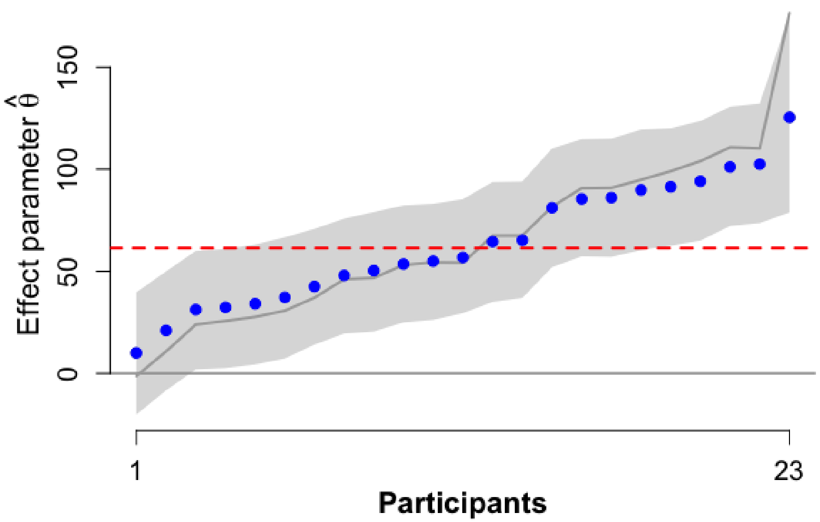
Bayes factor computations



E = encompassing approach

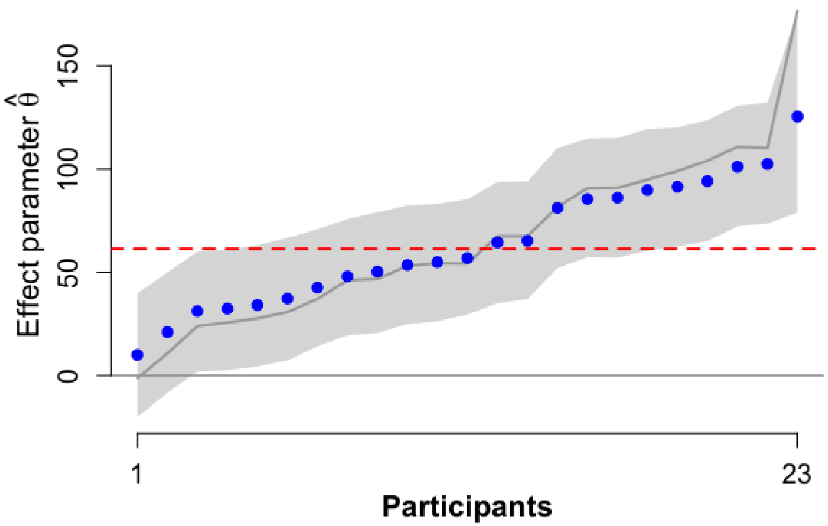
- Klugkist et al. (2005)
- $B_{+u} = \dfrac{P(\boldsymbol{\theta} > 0 \mid \boldsymbol{Y}, \mathcal{M}_u)}{P(\boldsymbol{\theta} > 0 \mid \mathcal{M}_u)}$
- probabilities computed as fraction of MCMC samples from unrestricted model that are **positive** for all individuals (both *a priori* and *a posteriori*)

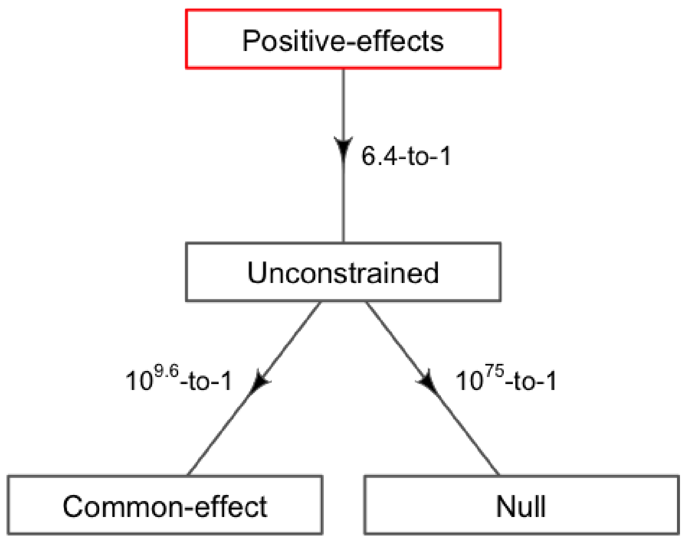
Results - Exp 1



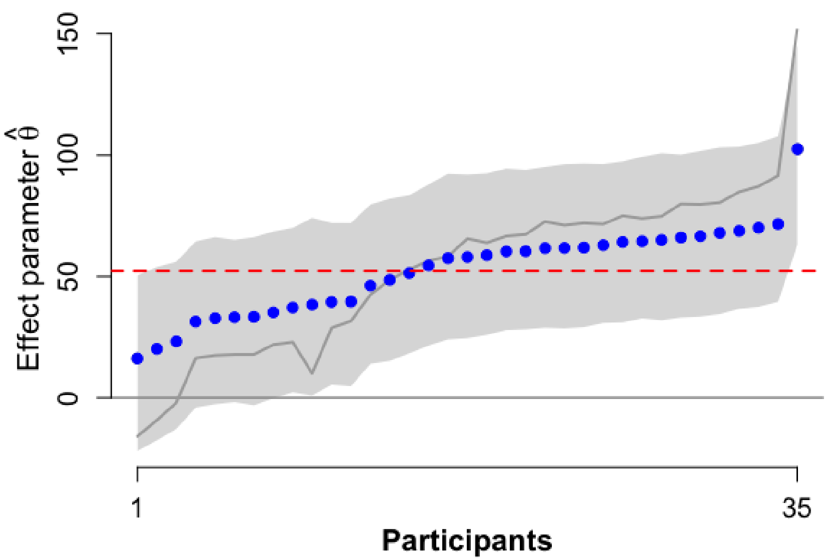
- Red line = estimated effect θ from $\mathcal{M}_1$
- Blue dots = individual effect estimates θ_i
- Gray line = estimates from mean differences d_i
- Gray area = 95% credible intervals

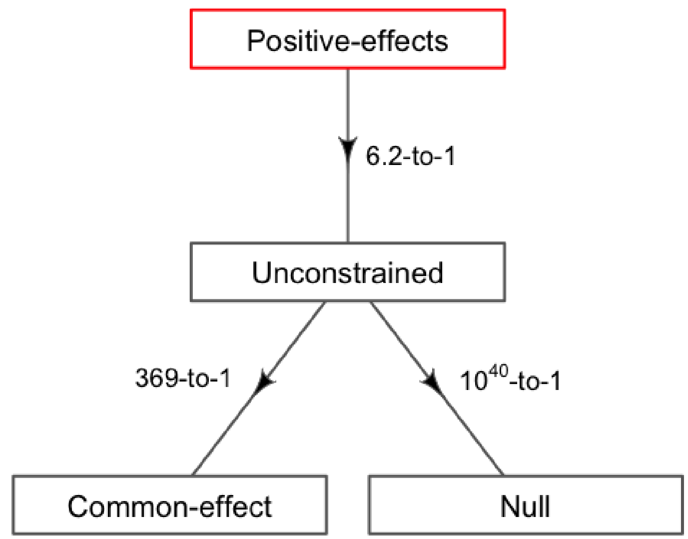
Results - Exp 1





Results - Exp 2





Sensitivity to prior specifications

Experiment 1:

r_ν	r_θ	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_+$	$\mathcal{M}_u$
$\frac{1}{6}$ (50 ms)	$\frac{1}{10}$ (30 ms)	5.6e-77	4.2e-11	*	0.16
$\frac{1}{12}$ (25 ms)	$\frac{1}{20}$ (15 ms)	1.8e-76	7.7e-11	*	0.16
$\frac{1}{12}$ (25 ms)	$\frac{1}{5}$ (60 ms)	1.9e-77	8.1e-12	*	0.05
$\frac{1}{3}$ (100 ms)	$\frac{1}{20}$ (15 ms)	1.9e-76	1.9e-10	*	0.39
$\frac{1}{3}$ (100 ms)	$\frac{1}{5}$ (60 ms)	3.0e-77	3.1e-11	*	0.16

Note: Bayes factors computed against the “winning” model, denoted by *

Sensitivity to prior specifications

Experiment 2:

r_ν	r_θ	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_+$	$\mathcal{M}_u$
$\frac{1}{6}$ (50 ms)	$\frac{1}{10}$ (30 ms)	4.3e-41	0.0004	*	0.17
$\frac{1}{12}$ (25 ms)	$\frac{1}{20}$ (15 ms)	1.2e-40	0.0007	*	0.16
$\frac{1}{12}$ (25 ms)	$\frac{1}{5}$ (60 ms)	2.6e-41	0.0002	*	0.06
$\frac{1}{3}$ (100 ms)	$\frac{1}{20}$ (15 ms)	1.4e-40	0.0017	*	0.42
$\frac{1}{3}$ (100 ms)	$\frac{1}{5}$ (60 ms)	4.1e-41	0.0005	*	0.19

Note: Bayes factors computed against the "winning" model, denoted by *

Summary points

Does everybody *exhibit the size congruity effect?*

- if *yes*, then SCE is obligatory, resistant to strategic control, ...
- if *no*, then SCE is complex, malleable, ...

Importantly, both answers have downstream consequences for processing architecture of numerical cognition

- what does this say about *early vs. late* interaction debate (e.g., Faulkenberry et al., 2016; Sobel et al., 2016; 2017; Faulkenberry, Vick, & Bowman, in press)

Some other benefits:

- Bayes factors easy to interpret

- hierarchical structure removes trial noise from individual estimates

- common effect (CE) model provides important self-check:
  - if CE model is best, is our design adequate to capture individual differences

- Might be good approach to disentangle competing theories of mental arithmetic
  - Does everyone exhibit size-by-format interaction?
  - Does everyone reflect fast counting in small addition problems?

Thank you!

- Thanks to Tarleton Office of Research and Innovation for funding!
- slides available at `github.com/tomfaulkenberry/talks`
- Twitter: `@tomfaulkenberry`
- Email: `faulkenberry@tarleton.edu`