A single-boundary accumulator model of response times in a mental arithmetic task

Thomas J. Faulkenberry

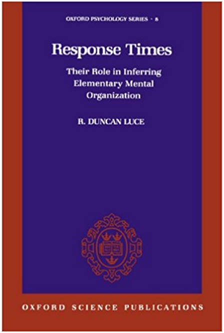
Tarleton State University

Outline

1. Importance of RTs for theorizing in mental arithmetic
2. Classical methods of measurement
3. Describing distributions via the Shifted Wald
4. Modeling some data in an arithmetic verification task
5. Extension to a hierarchical Bayesian model

Response times

Response times are a classic measure of cognitive processing



Preface

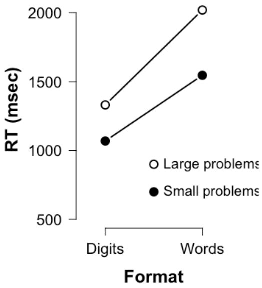
For almost as long as I have been doing research, response times have struck me as a fascinating, albeit tricky, source of information about how the mind is organized. Whenever I teach mathematical psychology or psychophysics, I include a dose of response times along with the repeated admonition that we surely do not understand a choice process very thoroughly until we can account for the time required for it to be carried out. When I came to Harvard in 1976 I offered, for the first time, a seminar-course on the subject (in style more a course, in size more a seminar), first with David Green and later alone. It was only then that I felt a need for a more systematic mastery of the field, and in academic 80–81 when I had a sabbatical leave, the Guggenheim Foundation agreed to support my self-education and the beginnings of this book.

Mental arithmetic

In mathematical cognition, RTs inform our understanding of the cognitive processes involved in mental arithmetic.

Mental arithmetic

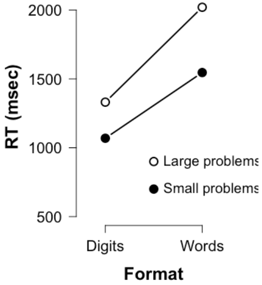
In mathematical cognition, RTs inform our understanding of the cognitive processes involved in mental arithmetic.



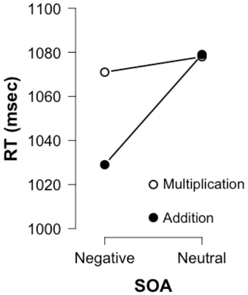
Campbell & Fugelsang (2001) - format-by-size interaction implies format directly affects calculation

Mental arithmetic

In mathematical cognition, RTs inform our understanding of the cognitive processes involved in mental arithmetic.



Campbell & Fugelsang (2001) - format-by-size interaction implies format directly affects calculation



Fayol & Thevenot (2012) - SOA effect for addition (but not multiplication) implies addition not solved via memory retrieval

Classical measurement

1. have people do a bunch of trials – measure RTs
2. find mean RT for each experimental condition / person
3. test for differences in means (e.g., ANOVA)

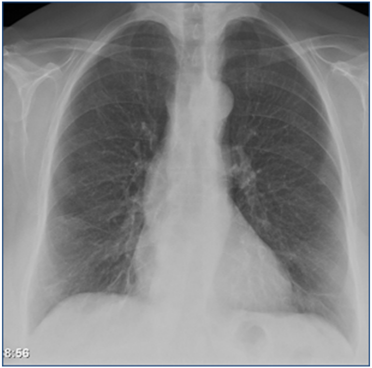
This method is lossy…it collapses each person's RT distribution to a single number

Measurement resolution

Low measurement resolution can result in not seeing what's there

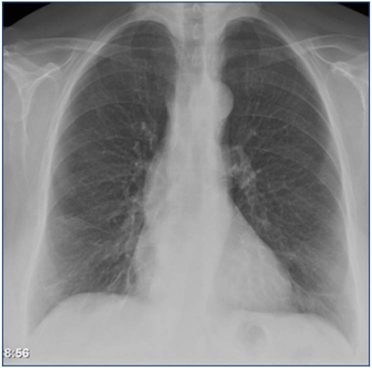
Measurement resolution

Low measurement resolution can result in not seeing what's there



Measurement resolution

Low measurement resolution can result in not seeing what's there

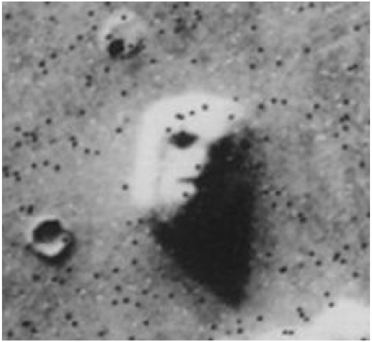


Measurement resolution

Low measurement resolution can result in seeing what's not there

Measurement resolution

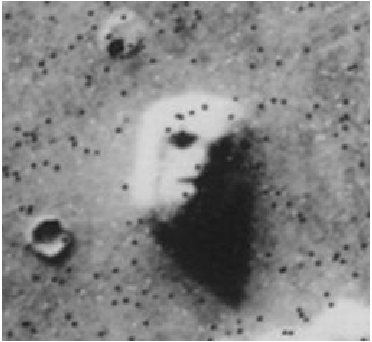
Low measurement resolution can result in seeing what's not there



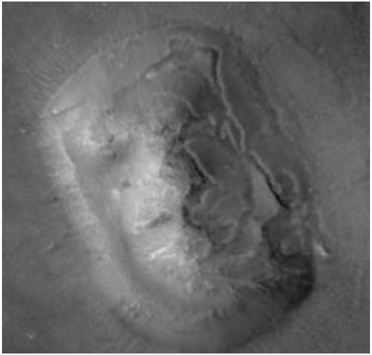
1976 - Viking I Orbiter

Measurement resolution

Low measurement resolution can result in seeing what's not there



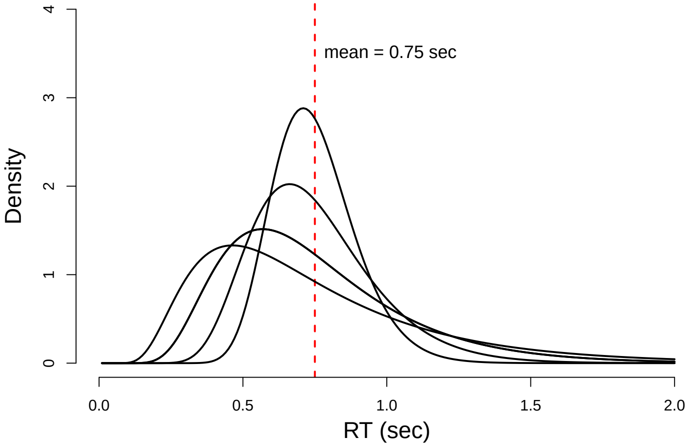
1976 - Viking I Orbiter



2006 - Mars Global Surveyor

RT distributions

The mean does not uniquely identify the distribution!



RT distributions

Need a measurement model that captures important characteristics of RT distributions

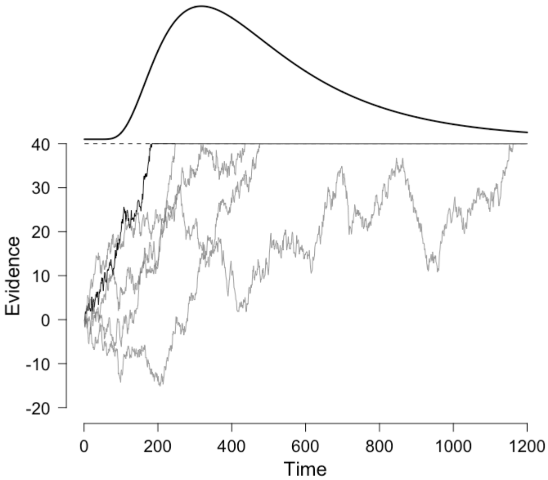
1. location
2. scale
3. shape

Some common models of RT

1. Weibull distribution (Rouder et al., 2005)
2. Lognormal distribution (van der Linden, 2006; Rouder & Province, submitted)
3. ex-Gaussian distribution (Campbell & Penner-Wilger, 2006; Heathcote et al., 1991)
4. Shifted Wald distribution (Heathcote, 2004; Anders et al., 2016)

Shifted Wald distribution

Density of hitting times for single-boundary diffusion process



Shifted Wald distribution

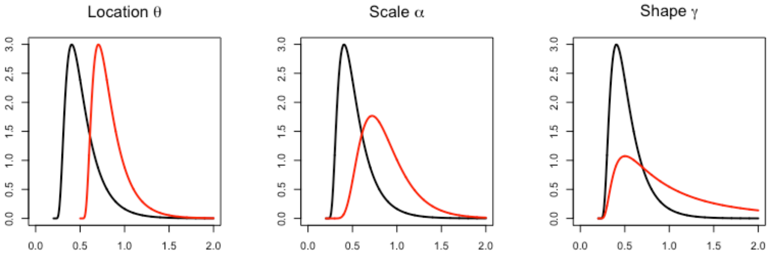
Common parameterization:

$$f(x) = \frac{\alpha}{\sqrt{2\pi(x-\theta)^3}} \exp\left(-\frac{(\alpha - \gamma(x-\theta)^2)}{2(x-\theta)}\right)$$

where:

- θ is a *location* parameter (nondecision time)
- α is a *scale* parameter (response caution)
- γ is a *shape* parameter (drift rate / evidence accumulation)

Shifted Wald distribution



- changes in θ affect onset of distribution
- changes in α affect deviation around the mode
- changes in γ affect mass in the tail

Fitting RT models

To fit a model to an RT distribution, we need to find an *optimal* set of parameters

Two methods:

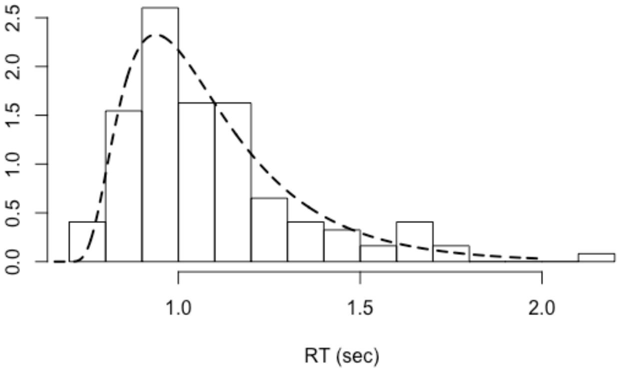
- maximum likelihood estimation
- hierarchical Bayesian model

Maximum likelihood estimation

Basic workflow:

1. collect data
2. decide on a model for the data (e.g., Weibull, lognormal, shifted Wald)
3. define a likelihood function based on the underlying model
4. find the parameter value(s) that maximize the likelihood function

Fitting a Shifted Wald model



RTs in mental arithmetic

Faulkenberry (2017) – 20 participants completed arithmetic verification task

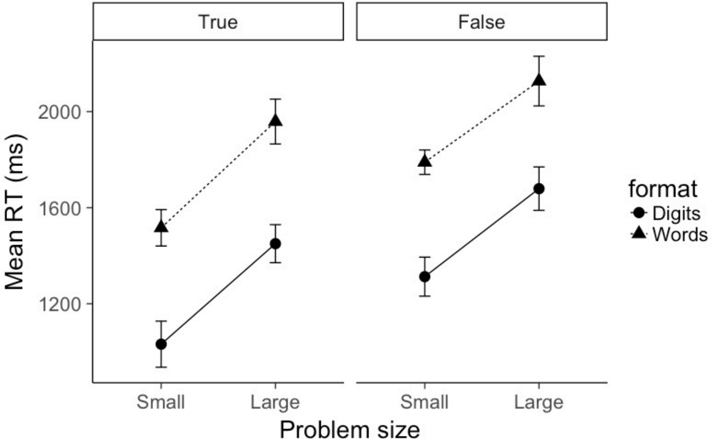
- format: digits versus words
- problem size: small versus large

$$2 + 3 = 7$$

$$\text{two} + \text{three} = \text{seven}$$

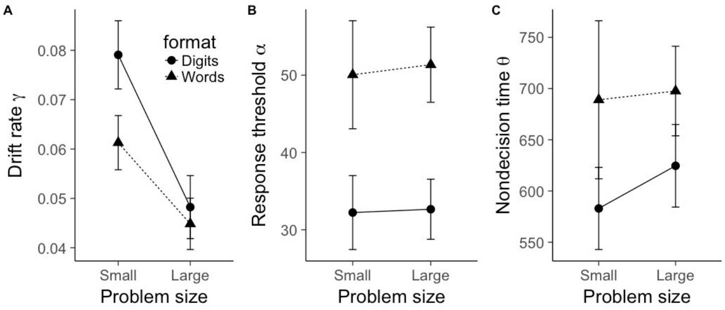
RTs in mental arithmetic

Model 1: mean RTs as a (linear) function of format and problem size



RTs in mental arithmetic

Model 2: fit RT distributions for each condition using shifted Wald model

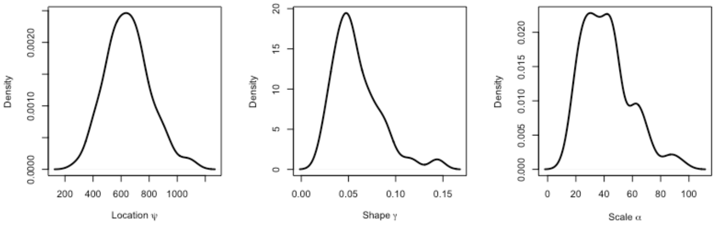


RTs in mental arithmetic

- Mean RTs did not reveal size-by-format interaction
- but, format did affect drift rate γ
- varying problem format has direct impact on calculation processes

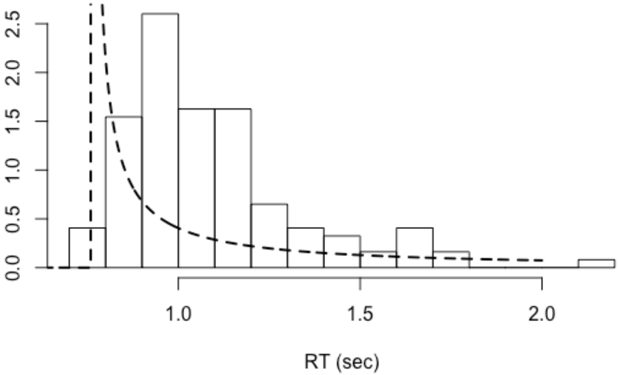
Problems with MLE approach

- parameters not normally distributed



Problems with MLE approach

- misfit due to "local minima"



Bayesian modeling

Basic workflow:

1. start with prior beliefs about parameter distribution
2. update prior to posterior distribution via Bayes Theorem

$$\text{posterior} \propto \text{likelihood} \cdot \text{prior}$$

Fitting Bayesian models

In practice, we estimate the posterior distribution

Basic recipe:

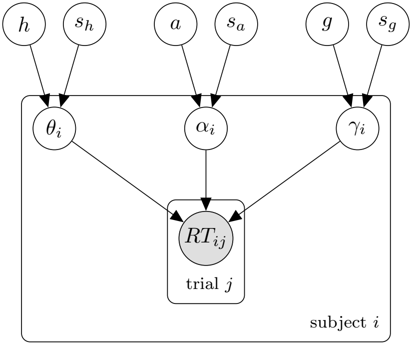
1. compute un-normalized posterior via Bayes theorem
2. sample from posterior (e.g., Markov chain Monte Carlo methods)
3. compute things from the samples

Bayesian advantages

1. posterior probability is often exactly what we want
2. tells us how data informs prior beliefs
3. ”easy” to build hierarchical models, allowing us to simultaneously model individual and group-level variation

Some new work

Hierarchical Bayesian model of response times in mental arithmetic



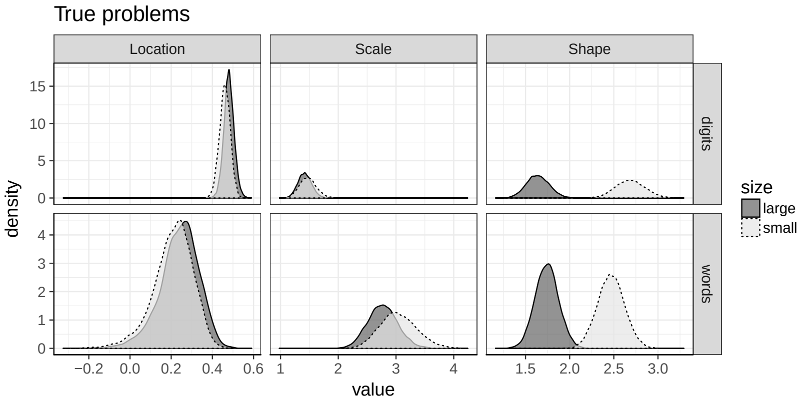
$$
\begin{aligned}
h &\sim \text{Gaussian}(0.61, 0.15^2) \\
s_h &\sim \text{Uniform}(0.01, 0.15) \\
a &\sim \text{Gamma}(8.86, 0.17) \\
s_a &\sim \text{Uniform}(0.01, \sqrt{0.256}) \\
g &\sim \text{Gamma}(10.32, 0.16) \\
s_g &\sim \text{Uniform}(0.01, \sqrt{0.264})
\end{aligned}
$$

$$
\begin{aligned}
\theta_i &\sim \text{Gaussian}(h, s_h^2) \\
\alpha_i &\sim \text{Gaussian}(a, s_a^2) \\
\gamma_i &\sim \text{Gaussian}(g, s_g^2)
\end{aligned}
$$

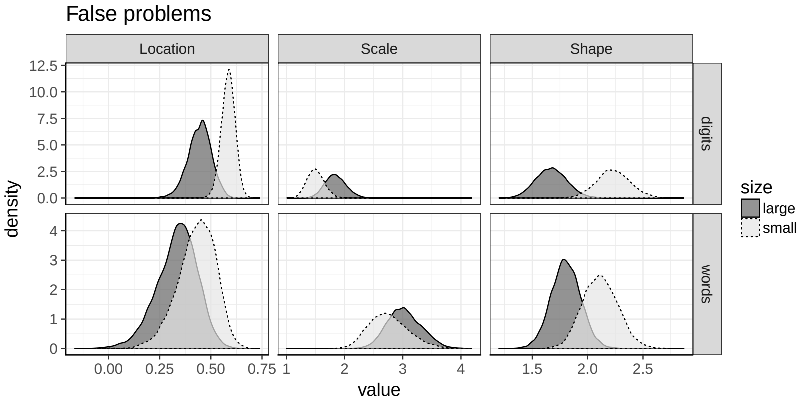
$$RT_{ij} \sim \text{ShiftedWald}(\theta_i, \alpha_i, \gamma_i)$$

based on Matzke & Wagenmakers (2009)

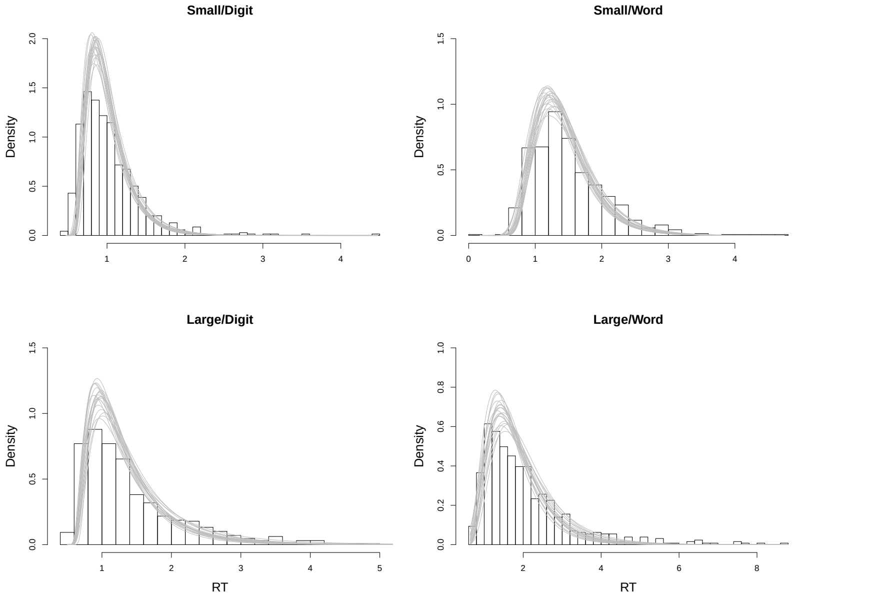
Posterior distributions



Posterior distributions



Posterior predictive checks



Stay tuned...

Next steps

- So far, each experimental condition is modeled separately
- next, need to build and compare models to estimate “effects” of problem size and format manipulations
- need to conduct parameter recovery simulations to test whether the model is robust to misspecification

Poster session tonight

Go see my student!

- Bowman, K. A., & Faulkenberry, T. J. – The dynamics of spatial-operational momentum in mental arithmetic (poster 3083)

Contact info

- Thomas J. Faulkenberry
- Department of Psychological Sciences
- Tarleton State University
- `faulkenberry@tarleton.edu`
- Twitter: `@tomfaulkenberry`