Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder

University of Missouri-Columbia

July, 2008
Collaborators

- Mike Pratte (Hire Him)
- Richard Morey (Too Late)
We have seen a plethora of signal detection and multinomial processing tree models

What I have to say today applies to them both.

I promised a process-dissociation (MPT) and a signal-detection application in my abstract.

I lied. Signal-detection application only.

Process Dissociation in Rouder et al., 2008, JEP:G.
We have seen a plethora of signal detection and multinomial processing tree models.
What I have to say today applies to them both.
- We have seen a plethora of signal detection and multinomial processing tree models.
- What I have to say today applies to them both.
- I promised a process-dissociation (MPT) and a signal-detection application in my abstract.
We have seen a plethora of signal detection and multinomial processing tree models.

What I have to say today applies to them both.

I promised a process-dissociation (MPT) and a signal-detection application in my abstract.

I lied. Signal-detection application only.
We have seen a plethora of signal detection and multinomial processing tree models.

What I have to say today applies to them both.

I promised a process-dissociation (MPT) and a signal-detection application in my abstract.

I lied. Signal-detection application only.

Process Dissociation in Rouder et al., 2008, JEP:G.
Hierarchical Models of Mnemonic Processes.
Input
Condition
A
B
C

Output

A
B
C
Hierarchical Models of Mnemonic Processes.
Input
Condition
A
B
C
Output
Input
s1, i1, cA
s1, i2, cA
s1, i3, cB
s1, i4, cB
s2, i1, cB
s2, i2, cA
s2, i3, cA
s2, i4, cB

Output
0
1
1
1
0
0
1
0
Input
s1, i1, cA
s1, i2, cA
s1, i3, cB
s1, i4, cB
s2, i1, cB
s2, i2, cA
s2, i3, cA
s2, i4, cB

Output
?

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Input

s1, i1, cA
s1, i2, cA
s1, i3, cB
s1, i4, cB
s2, i1, cB
s2, i2, cA
s2, i3, cA
s2, i4, cB

Output

m
n
o
p

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Input

s1, i1, cA
s1, i2, cA
s1, i3, cB
s1, i4, cB
s2, i1, cB
s2, i2, cA
s2, i3, cA
s2, i4, cB

Output
Hierarchical Models of Mnemonic Processes.
Two Views

Optimist: Aggregated data do indeed reflect underlying cognition:
- Little variation in transition time
- Learning really occurs gradually

Realist: We can learn about mixtures of cognitive processes.
- Confabulation of nuisance variation (people, items) with cognitive structure
- We cannot uncover certain processing invariances or structures

This problem is quite hard
Two Views

- **Optimist**: Aggregated data do indeed reflect underlying cognition:
  - Little variation in transition time
  - Learning really occurs gradually

- **Realist**: We can learn about mixtures of cognitive processes.
  - Confabulation of nuisance variation (people, items) with cognitive structure
  - We cannot uncover certain processing invariances or structures

This problem is quite hard
Two Views

- **Optimist**: Aggregated data do indeed reflect underlying cognition:
  - Little variation in transition time
  - Learning really occurs gradually

- **Realist**: We can learn about mixtures of cognitive processes.
  - Confabulation of nuisance variation (people, items) with cognitive structure
  - We cannot uncover certain processing invariances or structures
Two Views

**Optimist:** Aggregated data do indeed reflect underlying cognition:
- Little variation in transition time
- Learning really occurs gradually

**Realist:** We can learn about mixtures of cognitive processes.
- Confabulation of nuisance variation (people, items) with cognitive structure
- We cannot uncover certain processing invariances or structures

This problem is quite hard
Recognition Memory
We aggregate to construct rates.
We aggregate to construct rates.

There is an awareness on the effects of aggregating across people.
Recognition Memory

- We aggregate to construct rates.
- There is an awareness on the effects of aggregating across people.
- There is less awareness about aggregating across items.
We aggregate to construct rates.

There is an awareness on the effects of aggregating across people.

There is less awareness about aggregating across items.

If you use rates, you must aggregate across something.
Recognition Memory

Are you an optimist or a realist?

Optimist: Mnemonic structure accessible from rates.

Realist: Mnemonic structure is conflated with the distribution of items in the lexicon and people in the population.

Today: Optimist vs. realist view in signal detection.

How to have your cake and eat much of it to.

Jeffrey N. Rouder

University of Missouri-Columbia

Hierarchical Models of Mnemonic Processes.
Are you an optimist or a realist?

- Optimist: Mnemonic structure accessible from rates.
- Realist: Mnemonic structure is conflated with the distribution of items in the lexicon and people in the population.
Are you an optimist or a realist?

- Optimist: Mnemonic structure accessible from rates.
- Realist: Mnemonic structure is conflated with the distribution of items in the lexicon and people in the population.

Today: Optimist vs. realist view in signal detection.
Are you an optimist or a realist?
- Optimist: Mnemonic structure accessible from rates.
- Realist: Mnemonic structure is conflated with the distribution of items in the lexicon and people in the population.

Today: Optimist vs. realist view in signal detection.

How to have your cake and eat much of it to.
Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Hierarchical Models of Mnemonic Processes.
Hierarchical Models of Mnemonic Processes.
Hierarchical Models of Mnemonic Processes.
The Variance Benchmark Finding

\[ \sigma \approx 1.25 \]
Examples of Optimists (Who Aren’t Here)

Ben Murdock. Revised TODAM to account for $\sigma > 1$.

Andy Yonelinas. $\sigma > 1$ reflects mixture between signal-detection with $\sigma = 1$ and recollection.

Larry De Carlo. $\sigma > 1$ reflects mixture between signal-detection with $\sigma = 1$ and guessing.
Ben Murdock. Revised TODAM to account for $g \sigma > 1$. 
Examples of Optimists (Who Aren’t Here)

- Ben Murdock. Revised TODAM to account for $g \sigma > 1$.
- Andy Yonelinas. $\sigma > 1$ reflects mixture between signal-detection with $\sigma = 1$ and recollection.
Examples of Optimists (Who Aren’t Here)

- Ben Murdock. Revised TODAM to account for $g \sigma > 1$.
- Andy Yonelinas. $\sigma > 1$ reflects mixture between signal-detection with $\sigma = 1$ and recollection.
- Larry De Carlo. $\sigma > 1$ reflects mixture between signal-detection with $\sigma = 1$ and guessing.
Example of a Realist

On rare occasions, Wixted makes a realistic point. "The targets can be thought of as lures that have had memory strength added to them by virtue of their appearance on the study list. An equal-variance model would result if each item on the list had the exact same amount of strength added during study. However, if the amount of strength that is added differs across items, as it must, then both strength and variability would be added, and an unequal-variance model would apply."
Example of a Realist

- On rare occasions, Wixted makes a realistic point
Example of a Realist

- On rare occasions, Wixted makes a realistic point
- “The targets can be thought of as lures that have had memory strength added to them by virtue of their appearance on the study list. An equal-variance model would result if each item on the list had the exact same amount of strength added during study. However, if the amount of strength that is added differs across items, as it must, then both strength and variability would be added, and an unequal-variance model would apply.”
Possible Effect of Item Variability

- Half are Easy Items: $d' = 2.3$
- Half are Hard Items: $d' = .7$
- Overall: $d' = 1.5$
- Equal variance $\sigma = 1$
- Manipulate criterion through confidence ratings
- Recover $d' = 1.5, \sigma = 1$. 

Jeffrey N. Rouder
University of Missouri-Columbia
Possible Effect of Item Variability

Low criterion: $c = .4$

- Easy Hit: $\Phi(2.3 - .4) = .97$
- Easy FA: $\Phi(-.4) = .34$
- Hard Hit: $\Phi(.7 - .4) = .62$
- Hard FA: $\Phi(-.4) = .34$
- Average Hit: $(.97 + .62)/2 = .80$
- Average FA: .34
Possible Effect of Item Variability

High criterion: $c = 1.1$

- Easy Hit: $\Phi(2.3 - 1.1) = .86$
- Easy FA: $\Phi(-1.1) = .14$
- Hard Hit: $\Phi(.7 - 1.1) = .36$
- Hard FA: $\Phi(-1.1) = .14$
- Average Hit: $(.96 + .36)/2 = .61$
- Average FA: .14

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Possible Effect of Item Variability

Hit Rate vs. False-Alarm Rate

\( \sigma = 1.24 \)
It is Important to Separate Item, People, and Process Variability

All mnemonic theories have core assumptions about process variability that are independent of the distribution of participant abilities or item effects.
Aggregation in MPT
Aggregation in MPT

- Asymptotic distortion of parameter estimates
Aggregation in MPT

- Asymptotic distortion of parameter estimates
- Inflated Type I Error Rates
Aggregation in MPT

- Asymptotic distortion of parameter estimates
- Inflated Type I Error Rates
- Plays havoc with selective-influence validation tests
What is the value of $\sigma$ from core cognitive processes.
What is the value of $\sigma$ from core cognitive processes.

What is the value of $\sigma$ if we could observe replicates of people-by-item combinations.
For all-subject-by-item combinations, $\sigma = 1$. Estimates of $\sigma > 1$ because of the effects of item/participant variation on aggregated rates. Parsimony: Effect of study, item, and people is to shift distributions of strength.

Join me in then working conjecture?

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
For all-subject-by-item combinations, $\sigma = 1$. 
For all-subject-by-item combinations, $\sigma = 1$.

Estimates of $\sigma > 1$ because of the effects of item/participant variation on aggregated rates.
For all-subject-by-item combinations, $\sigma = 1$.

Estimates of $\sigma > 1$ because of the effects of item/participant variation on aggregated rates.

Parsimony: Effect of study, item, and people is to shift distributions of strength.
For all-subject-by-item combinations, $\sigma = 1$.

Estimates of $\sigma > 1$ because of the effects of item/participant variation on aggregated rates.

Parsimony: Effect of study, item, and people is to shift distributions of strength.

John Dunn’s single-factor model
Working Conjecture

- For all-subject-by-item combinations, $\sigma = 1$.
- Estimates of $\sigma > 1$ because of the effects of item/participant variation on aggregated rates.
- Parsimony: Effect of study, item, and people is to shift distributions of strength.
- John Dunn’s single-factor model
- Join me in then working conjecture?
Hierarchical model to account for process, person, and item variability simultaneously.
Can We Separate Process, Item, and Person Variability

- Yes We Can
Can We Separate Process, Item, and Person Variability

- Yes We Can
- Hierarchical model to account for process, person, and item variability simultaneously.
Hierarchical Models of Mnemonic Processes.
Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder
University of Missouri-Columbia
Bias Effects as Correlated Shifts

Memory Strength

Density

Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder

University of Missouri-Columbia
Bias Effects as Correlated Shifts

Density

Memory Strength

Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder
University of Missouri-Columbia
Bias Effects as Correlated Shifts

Mirror Effects as Neg. Cor. Shifts

Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder
University of Missouri-Columbia
If We Had Person-by-Item Replicates

- Person $i$
- Item $j$
- Response to studied item $y_{ij}^{(s)} = 1, \ldots, K$
- Response to novel item $y_{ij}^{(n)} = 1, \ldots, K$

$$Pr(y_{ij}^{(s)} = k) = \text{Area}(d_{ij}^{(s)}, \sigma, c_{ik})$$

$$Pr(y_{ij}^{(n)} = k) = \text{Area}(d_{ij}^{(n)}, 1, c_{ik})$$
Without Replicates

\[
d_{ij}^{(s)} = \text{Grand Mean}^{(s)} + \text{Person}_{i}^{(s)} + \text{Item}_{j}^{(s)}
\]

\[
d_{ij}^{(n)} = \text{Grand Mean}^{(n)} + \text{Person}_{i}^{(n)} + \text{Item}_{j}^{(n)}
\]
Without Replicates

\[
\begin{align*}
    d_{ij}^{(s)} &= \text{Grand Mean}^{(s)} + \text{Person}_i^{(s)} + \text{Item}_j^{(s)} \\
    d_{ij}^{(n)} &= \text{Grand Mean}^{(n)} + \text{Person}_i^{(n)} + \text{Item}_j^{(n)}
\end{align*}
\]

- Person and item effects are zero-centered normally-distributed random effects.
Without Replicates

\[
d_{ij}^{(s)} = \text{Grand Mean}^{(s)} + \text{Person}_i^{(s)} + \text{Item}_j^{(s)}
\]

\[
d_{ij}^{(n)} = \text{Grand Mean}^{(n)} + \text{Person}_i^{(n)} + \text{Item}_j^{(n)}
\]

- Person and item effects are zero-centered normally-distributed random effects.
- Priors are standard
Without Replicates

\[
d^{(s)}_{ij} = \text{Grand Mean}^{(s)} + \text{Person}_{i}^{(s)} + \text{Item}_{j}^{(s)}
\]

\[
d^{(n)}_{ij} = \text{Grand Mean}^{(n)} + \text{Person}_{i}^{(n)} + \text{Item}_{j}^{(n)}
\]

- Person and item effects are zero-centered normally-distributed random effects.
- Priors are standard
- Tad optimistic about a lack of participant-by-item interactions.
Without Replicates

\[ d_{ij}^{(s)} = \text{Grand Mean}^{(s)} + \text{Person}_{i}^{(s)} + \text{Item}_{j}^{(s)} \]
\[ d_{ij}^{(n)} = \text{Grand Mean}^{(n)} + \text{Person}_{i}^{(n)} + \text{Item}_{j}^{(n)} \]

- Person and item effects are zero-centered normally-distributed random effects.
- Priors are standard
- Tad optimistic about a lack of participant-by-item interactions.
- Covariates are easily added (e.g., lag)
Hierarchical Signal Detection

- Allows for accurate measurement of process parameters \((\sigma)\) without conflation from item and participant effects.
- Benchmarked through simulation.
Hierarchical Signal Detection

Hierarchical Models of Mnemonic Processes.

Jeffrey N. Rouder
University of Missouri-Columbia
Measuring $\sigma$

Experiment

- Near replication of Glanzer et al., 1999.
- 90 people observed 480 items at test.
- 240 items studied in a single list (2 seconds per item)
Measuring $\sigma$

Experiment

- Near replication of Glanzer et al., 1999.
- 90 people observed 480 items at test.
- 240 items studied in a single list (2 seconds per item)
- Overall $d' \approx 1.5$
Methods of Analysis

1. Double aggregation (across both participants and items)
2. Single aggregation (across only one at a time)
3. Hierarchical Model
Conventional Analysis

\[ d' = 1.23 \]
\[ \sigma = 1.28 \]
Measuring $\sigma$

Double Agg. Item Agg. People Agg. Hierarchical

$\sigma$

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Mnemonic Structure (Average item & person)

Latent Strength

Density

Hierarchical Models of Mnemonic Processes.
Conjecture is Wrong

New Benchmark: $\sigma \approx 1.4$. 
Unequal-Variance Normal Model Drawbacks

▶ Stochastic Indominance (De Carlo)
▶ Too flexible (Lockhart & Murdock)
▶ Effect of study in two loci: shift ($d'$) and scale ($\sigma$)

We should be placing hierarchical models on $\sigma$ (e.g., $\sigma_{ij} = \sigma_0 \theta_i \eta_j$) in addition to those on $d$ ($n$) and $d$ ($s$). Too complex.

▶ Search for a more parsimonious single-factor model.
Unequal-Variance Normal Model Drawbacks

- Stochastic Indominance (De Carlo)

We should be placing hierarchical models on $\sigma$ (e.g., $\sigma_{ij} = \sigma_0 \theta_i \eta_j$) in addition to those on $d$ ($n$) and $d$ ($s$). Too complex.

Search for a more parsimonious single-factor model.
Unequal-Variance Normal Model Drawbacks

- Stochastic Indominance (De Carlo)
- Too flexible (Lockhart & Murdock)
Unequal-Variance Normal Model Drawbacks

- Stochastic Indominance (De Carlo)
- Too flexible (Lockhart & Murdock)
- Effect of study in two loci: shift ($d'$) and scale ($\sigma$)

We should be placing hierarchical models on $\sigma$ (e.g., $\sigma_{ij} = \sigma_0 \theta_i \eta_j$) in addition to those on $d'$ ($n$) and $d'$ ($s$). Too complex.

Search for a more parsimonious single-factor model.
Unequal-Variance Normal Model Drawbacks

- Stochastic Indominance (De Carlo)
- Too flexible (Lockhart & Murdock)
- Effect of study in two loci: shift ($d'$) and scale ($\sigma$)
- We should be placing hierarchical models on $\sigma$ (e.g., $\sigma_{ij} = \sigma_0 \theta_i \eta_j$) in addition to those on $d^{(n)}$ and $d^{(s)}$. Too complex.
Unequal-Variance Normal Model Drawbacks

- Stochastic Indominance (De Carlo)
- Too flexible (Lockhart & Murdock)
- Effect of study in two loci: shift ($d'$) and scale ($\sigma$)
- We should be placing hierarchical models on $\sigma$ (e.g., $\sigma_{ij} = \sigma_0 \theta_i \eta_j$) in addition to those on $d^{(n)}$ and $d^{(s)}$. Too complex.
- Search for a more parsimonious single-factor model.
From Aggregates: $\sigma$ Increases with $d'$
Jeff & Mike’s Gamma Model

Latent Strength density

Jeffrey N. Rouder
University of Missouri-Columbia

Hierarchical Models of Mnemonic Processes.
Gamma Model

- The gamma distribution: scale ($\theta$) & shape.

- Fix shape = 2. Common in electronics, hydrology.

- Mean & standard deviation are both proportional to scale.

- Study affects one factor: scale

- People and item effects are in scale too

- As parsimonious as equal-variance, without equal variances

- Correct ROC predictions

- Solves John Dunn's variance problem
The gamma distribution: scale ($\theta$) & shape.

Fix shape = 2. Common in electronics, hydrology.
The gamma distribution: scale ($\theta$) & shape.

- Fix shape $= 2$. Common in electronics, hydrology.
- Mean & standard deviation are both proportional to scale.
The gamma distribution: scale ($\theta$) & shape.

- Fix shape = 2. Common in electronics, hydrology.
- Mean & standard deviation are both proportional to scale.
- Study affects one factor: scale
Gamma Model

- The gamma distribution: scale ($\theta$) & shape.
- Fix shape = 2. Common in electronics, hydrology.
- Mean & standard deviation are both proportional to scale.
- Study affects one factor: scale
- People and item effects are in scale too
The gamma distribution: scale ($\theta$) & shape.

- Fix shape = 2. Common in electronics, hydrology.
- Mean & standard deviation are both proportional to scale.
- Study affects one factor: scale
- People and item effects are in scale too
- As parsimonious as equal-variance, without equal variances
The gamma distribution: scale ($\theta$) & shape.

- Fix shape = 2. Common in electronics, hydrology.
- Mean & standard deviation are both proportional to scale.
- Study affects one factor: scale
- People and item effects are in scale too
- As parsimonious as equal-variance, without equal variances
- Correct ROC predictions
The gamma distribution: scale ($\theta$) & shape.
Fix shape = 2. Common in electronics, hydrology.
Mean & standard deviation are both proportional to scale.
Study affects one factor: scale
People and item effects are in scale too
As parsimonious as equal-variance, without equal variances
Correct ROC predictions
Solves John Dunn’s variance problem
Hierarchical Gamma Model

\[
Pr(y_{ij}^{(s)} = k) = \text{Area}(\theta_{ij}^{(s)}, c_{ik})
\]

\[
Pr(y_{ij}^{(n)} = k) = \text{Area}(\theta_{ij}^{(n)}, c_{ik})
\]

where

\[
\theta_{ij}^{(s)} = \text{Grand Mean}^{(s)} \times \text{Person}_i^{(s)} \times \text{Item}_j^{(s)}
\]

\[
\theta_{ij}^{(n)} = \text{Grand Mean}^{(n)} \times \text{Person}_i^{(n)} \times \text{Item}_j^{(n)}
\]

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Mnemonic Structure for Ave. Item & Person

Hierarchical Models of Mnemonic Processes.
Multiplier Effects

People Effects

Item Effects

Jeffrey N. Rouder
University of Missouri-Columbia
Hierarchical Models of Mnemonic Processes.
Correlated Effects

People Scales

Item Scales
A Very Convenient Approach

- $\theta^{(k)} \in [0, 1]$ for $k = 1, \ldots, K$
- $\theta^{(k)} = \Phi(\eta^{(k)})$
- $\eta_{ij}^{(k)} = \alpha_i^{(k)} + \beta_j^{(k)}$
Conclusion

1. Cognitive structure is separate from item and participant effects.
2. Hierarchical models may be used to separate people, item, and process variation, even without replicates.
3. When item and participant effects are accounted, $\sigma \approx 1$.
4. Model: the effect of study is to scale positive-going strength distributions.
5. People display larger response-bias effects than mirror effects; items display larger mirror effects than response-bias effects.
Conclusion

1. Cognitive structure is separate from item and participant effects.
Conclusion

1. Cognitive structure is separate from item and participant effects.

2. Hierarchical models may be used to separate people, item, and process variation, even without replicates.
Conclusion

1. Cognitive structure is separate from item and participant effects.

2. Hierarchical models may be used to separate people, item, and process variation, even without replicates.

3. When item and participants effects are accounted, $\sigma \approx 1.4$. 
Conclusion

1. Cognitive structure is separate from item and participant effects.
2. Hierarchical models may be used to separate people, item, and process variation, even without replicates.
3. When item and participants effects are accounted, $\sigma \approx 1.4$.
4. Model: the effect of study is to scale positive-going strength distributions.
1. Cognitive structure is separate from item and participant effects.

2. Hierarchical models may be used to separate people, item, and process variation, even without replicates.

3. When item and participants effects are accounted, $\sigma \approx 1.4$.

4. Model: the effect of study is to scale positive-going strength distributions.

5. People display larger response-bias effects than mirror effects; items display larger mirror effects than response-bias effects.
How May I Help You (Be a Realist)?

If you could easily implement a hierarchical model to isolate cognitive processes, would you?

▶

If so, how do we develop this methodological infrastructure?
If you could easily implement a hierarchical model to isolate cognitive process, would you?
How May I Help You (Be a Realist)?

- If you could easily implement a hierarchical model to isolate cognitive process, would you?
- If so, how do we develop this methodological infrastructure?