Understanding Changes in Episodic Memory Impairment Using Batchelder's Multinomial Processing Tree Model

Jason R. Bock [2] William R. Shankle [2] Isaiah Cushman [1] Michael D. Lee [1]

March 21, 2019

[1]University of California Irvine

[2]Medical Care Corporation

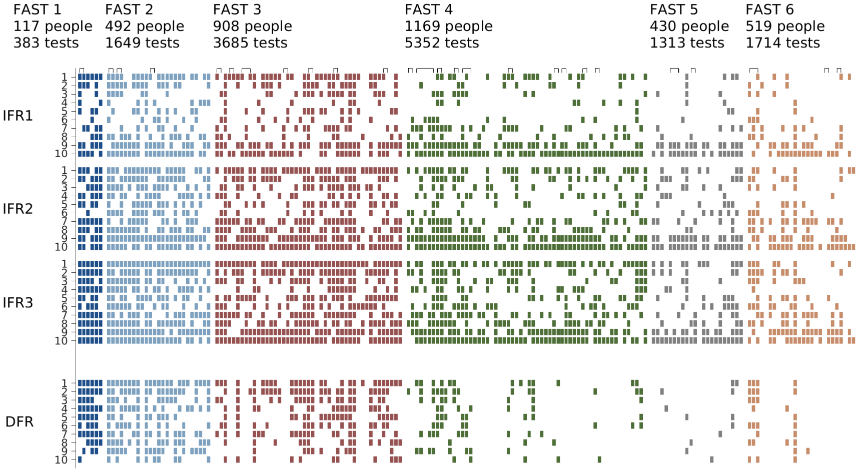
Bill Batchelder (1940–2018)

Shankle Clinic Data

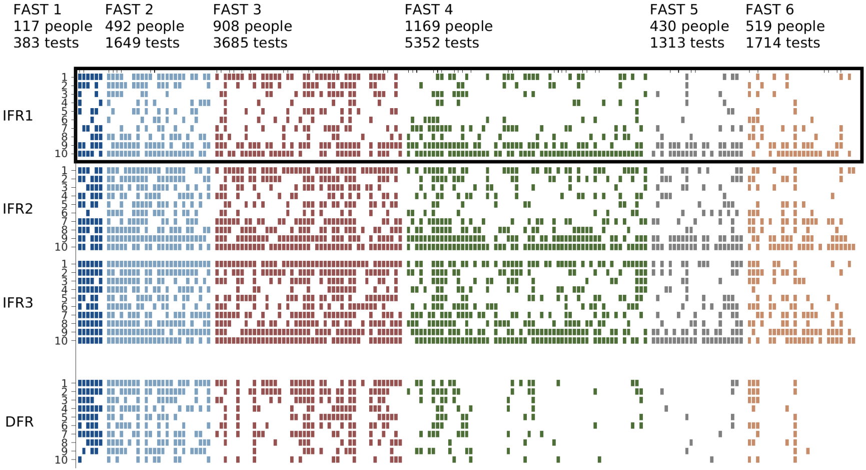
- Patients from a cognitive disorders clinic, given standard MCI screen assessment of memory
  - Total of 3635 patients doing a total of 14,096 assessments
  - A few patients do many assessments but most do only a few
- We focus on the free recall tasks
  - Three Immediate Free Recalls (IFR1, IFR2, IFR3) of the same list of 10 semantically-controlled words presented in the same order
  - A later expected Delayed Free Recall (DFR)

Stage	Name	Patients	Assessments
1	Normal aging	117	383
2	Possible mild cognitive impairment	492	1649
3	Mild cognitive impairment	908	3685
4	Mild dementia	1169	5352
5	Moderate dementia	430	1313
6	Moderately severe dementia	519	1714

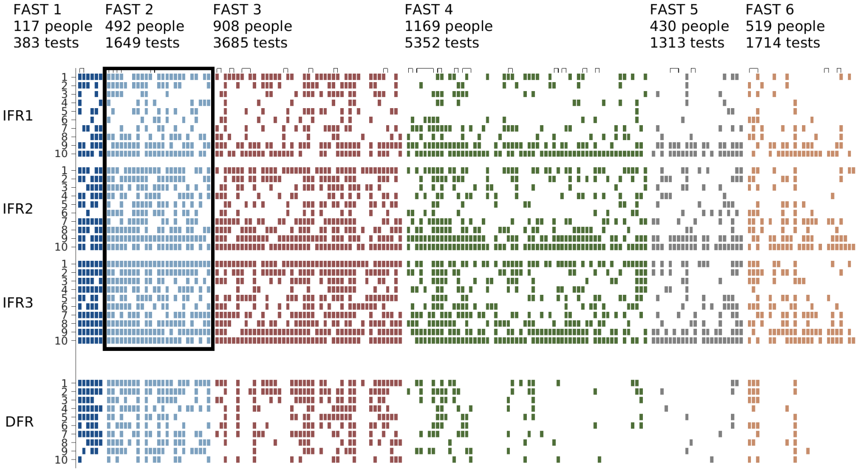
Visualization of 5% of Data



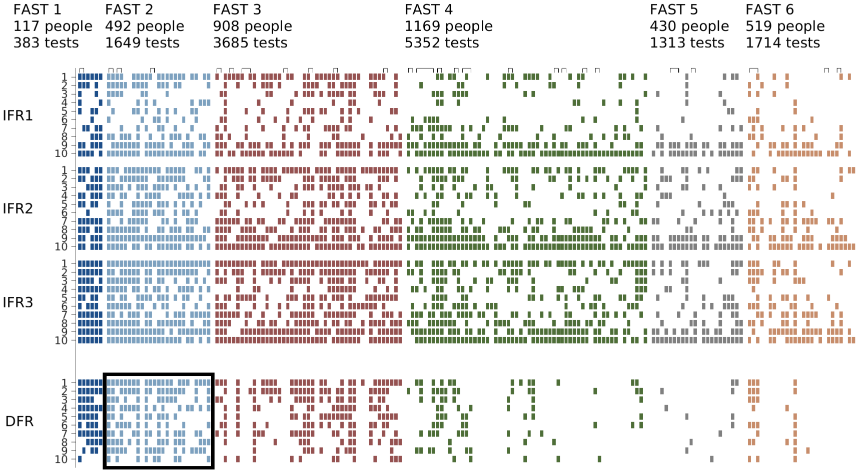
Worse Recall with Increasing Impairment



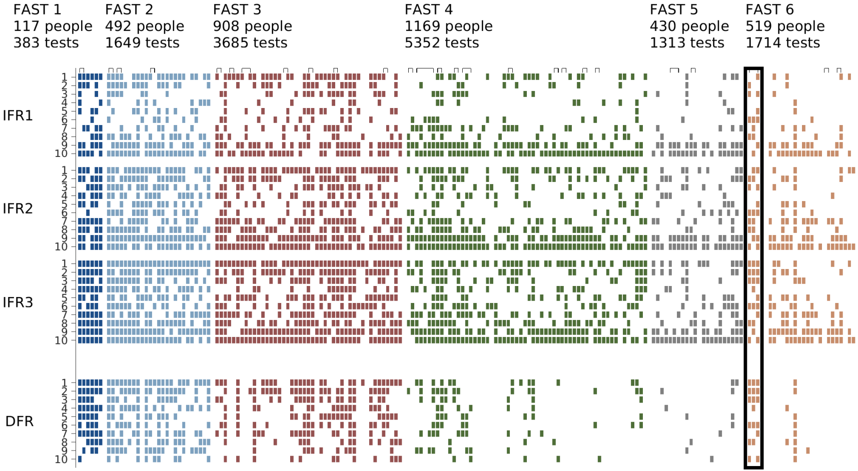
Serial Position Curves and Learning



No Recency Effect after Delay

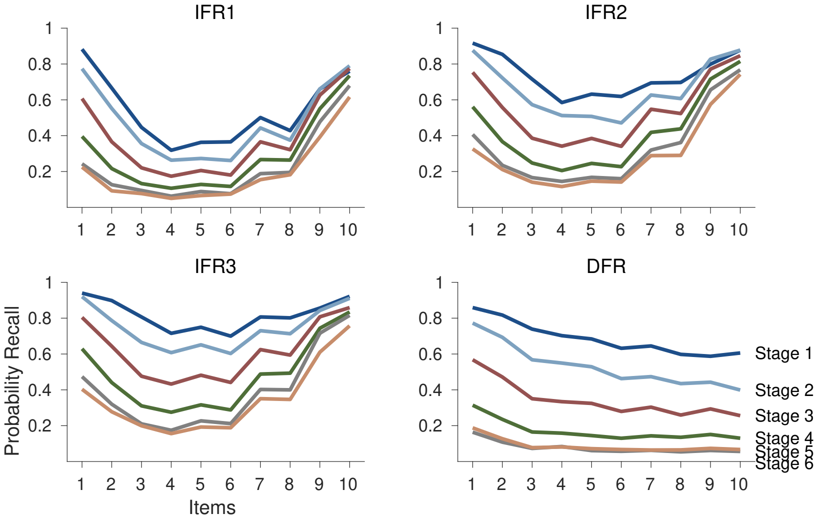


Individual Differences



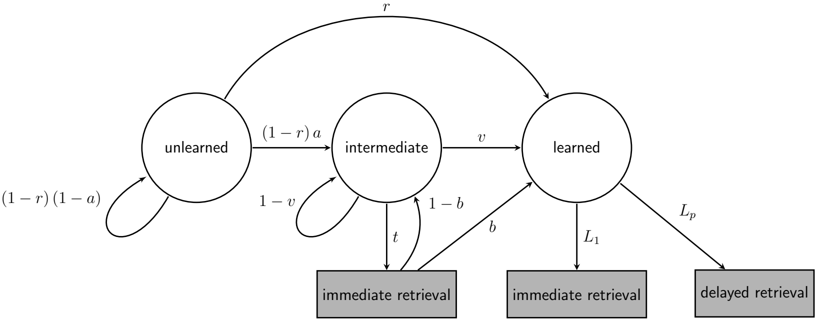
Serial Position Curves

- Free recall shows standard serial position curves
  - Learning over trials, but worsening performance with impairment



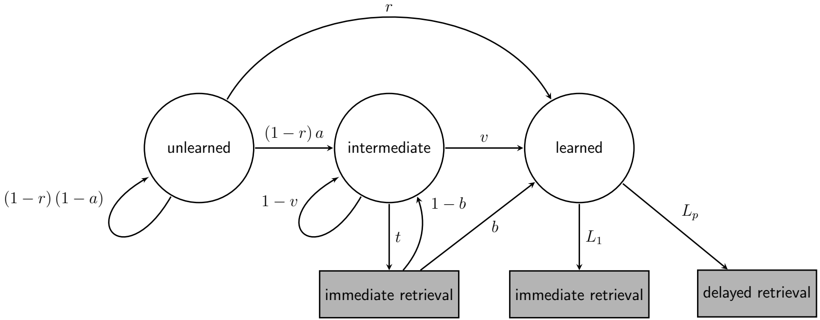
Multinomial Processing Tree Model

- Alexander, Satalich, Shankle, and Batchelder (2016) propose a MPT model of the retrieval of an item over a sequence of immediate and delayed free recall tasks
  - Key innovation is the assumption of unlearned, intermediate (partially-learned), and learned states for an item over testing



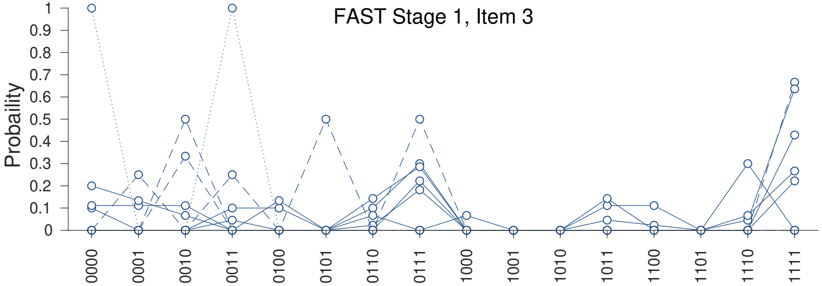
Application to MCI Screen Tasks

$$\theta_{0000} = \overbrace{r(1-L_1)(1-L_1)(1-L_1)(1-L_p)}^{\text{learned immediately but never retrieve}} + (1-r)a(1-t)v(1-L_1)(1-L_1)(1-L_p) + (1-r)(1-a)r(1-L_1)(1-L_1)(1-L_p) + (1-r)a(1-t)(1-v)(1-t)v(1-L_1)(1-L_p) + (1-r)(1-a)(1-r)a(1-t)v(1-L_1)(1-L_p) + (1-r)(1-a)(1-r)(1-a)r(1-L_1)(1-L_p) + (1-r)a(1-t)(1-v)(1-t)(1-v)(1-t) + (1-r)(1-a)(1-r)a(1-t)(1-v)(1-t) + (1-r)(1-a)(1-r)(1-a)(1-r)a(1-t) + \underbrace{(1-r)(1-a)(1-r)(1-a)(1-r)(1-a)}_{\text{always unlearned}}$$



16-Tuple Representation of Data

- Each item is either recalled or not recalled on each of the four recall tasks, giving $2^4 = 16$ possible outcomes
  - 1111 means the item was recalled every time
  - 1110 means the item was recalled for the first three immediate free recalls, but not the delayed free recall
  - ...
  - 0000 means the item was never recalled
- We represent behavioral data as counts y_{ij} of the jth of the 16-tuple patterns for the ith person over their n_i assessments

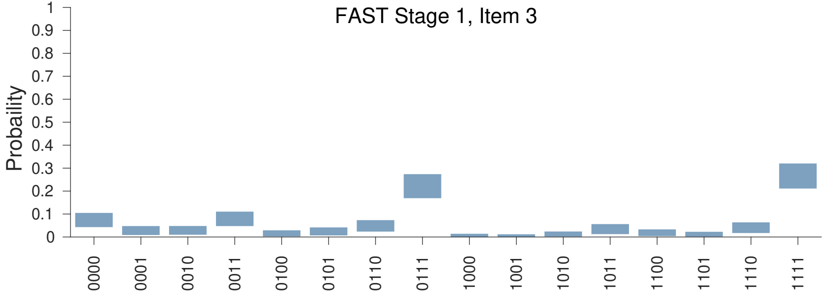


Saturated Model

- For posterior predictive checks of the descriptive adequacy of substantive models, we characterize the data by a saturated model

$$\begin{aligned} y_{ij} &\sim \text{Multinomial}\left(\boldsymbol{\theta}_{ij}, n_i\right) \\ \boldsymbol{\theta}_{ij} &\sim \text{Dirichlet}\left(\boldsymbol{\alpha}_j\right) \\ \alpha_{jk} &\sim \text{Gamma}\left(2, 1\right), \end{aligned}$$

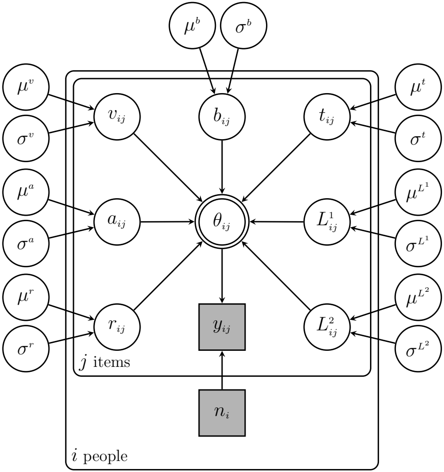
and the distribution of $\theta_j^{\text{pred}} \sim \text{Dirichlet}\left(\boldsymbol{\alpha}_j\right)$ quantifies the uncertainty of the probability of the jth tuple occuring for any person



Fixed Item Model

Fixed Item Model

Assume individual differences from a truncated Gaussian for each parameter, but no item differences in parameters



$$
\begin{aligned}
\mu^{\cdot} &\sim \text{Uniform}(0,1)\\
\sigma^{\cdot} &\sim \text{Uniform}(0,1)\\
\cdot_{ij} &\sim \text{TruncatedGaussian}_{(0,1)}\left(\mu^{\cdot}, 1/\sigma^{\cdot 2}\right)\\
\theta_{ij} &= \text{Batchelder}\left(a_{ij}, b_{ij}, r_{ij}, v_{ij}, t_{ij}, L^1_{ij}, L^2_{ij}\right)\\
y_{ij} &\sim \text{Multinomial}\left(\boldsymbol{\theta}_i, n_i\right)
\end{aligned}
$$

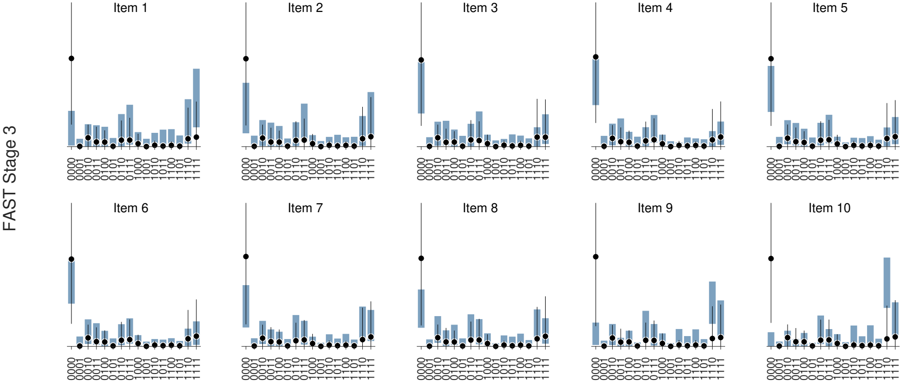
JAGS Implementation

JAGS script makes use of an user-added function `Batchelder` that returns the 16-tuple probabilities given a set of MPT parameters

```
model{
 for (j in 1:nItems){
   for (i in 1:nPeople){
    # data
    y[i,j,1:nPatterns]~dmulti(theta[i,j,1:nPatterns],nAssessments[i])
    # model
    theta[i,j,1:nPatterns]=Batchelder(a[i,j],b[i,j],r[i,j],v[i,j],t[i,j],L1[i,j],L2[i,j])
    # parameters
    a[i,j]~dnorm(mua,1/sigmaa^2)T(0,1)
    b[i,j]~dnorm(mub,1/sigmab^2)T(0,1)
    ...
   }
  }
 #  priors
 mua~dunif(0,1)
 sigmaa~dunif(0,1)
 mub~dunif(0,1)
 sigmab~dunif(0,1)
 ...
}
```

Failure of Descriptive Adequacy

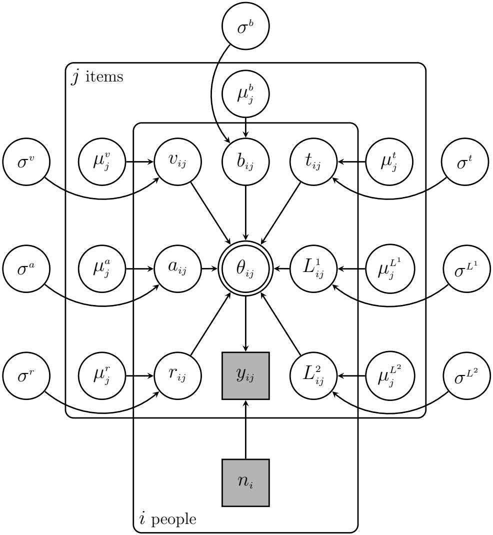
- A posterior predictive comparison of the distribution of observed proportions and the distribution over individual differences
  - The model cannot describe the data because different items have very different recall patterns
  - Example below is for FAST stage 3, but all stages have the same property



Independent Item Model

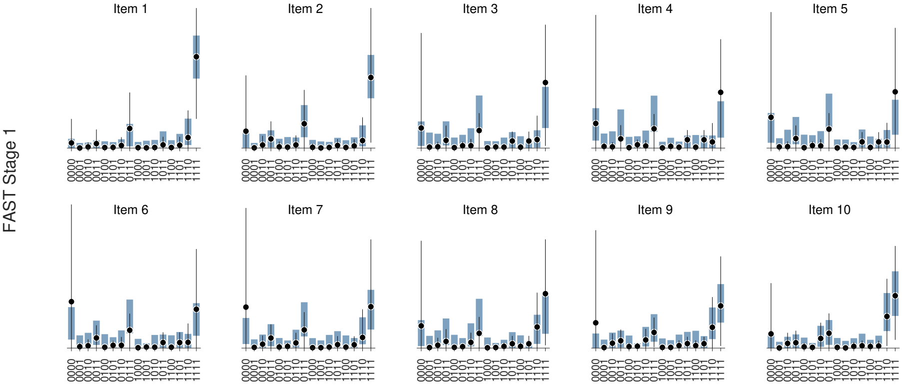
Independent Item Model

Assume individual differences from a truncated Gaussian for each parameter, and now allow independent parameters for each item position



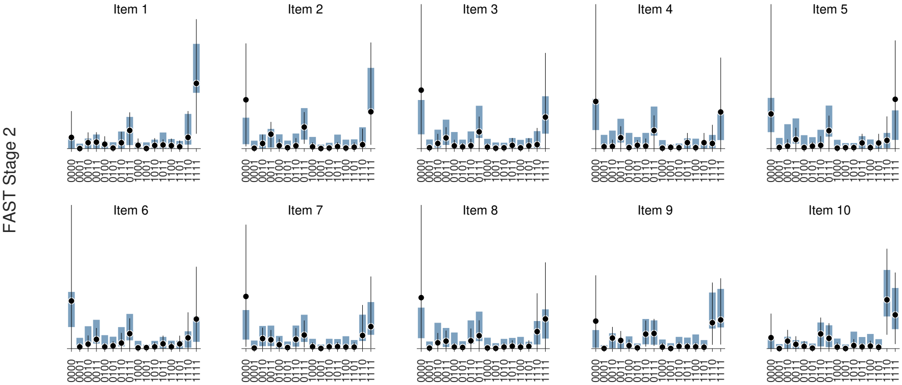
Descriptive Adequacy for Stage 1

- Allowing for each item to have its own parameterization achieves descriptive adequacy



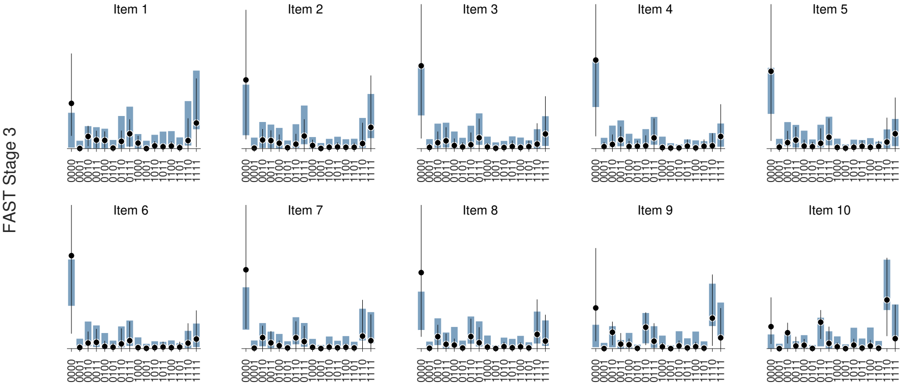
Descriptive Adequacy for Stage 2

- Allowing for each item to have its own parameterization achieves descriptive adequacy



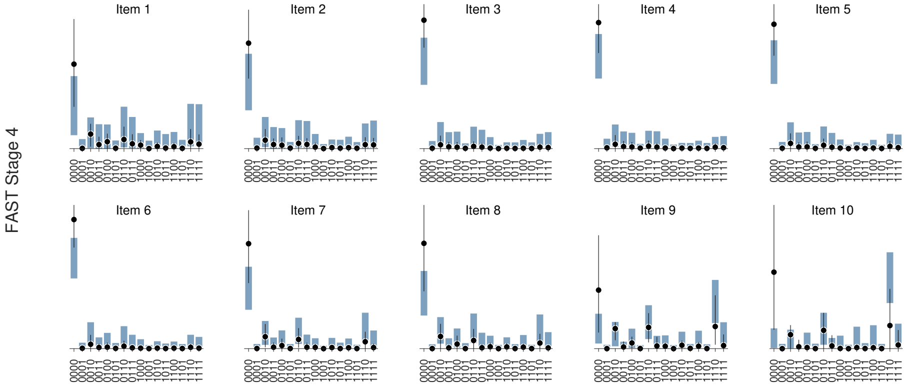
Descriptive Adequacy for Stage 3

- Allowing for each item to have its own parameterization achieves descriptive adequacy



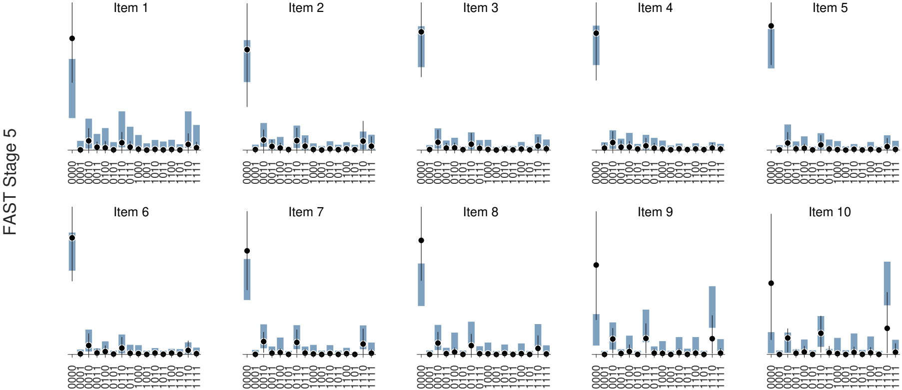
Descriptive Adequacy for Stage 4

- Allowing for each item to have its own parameterization achieves descriptive adequacy



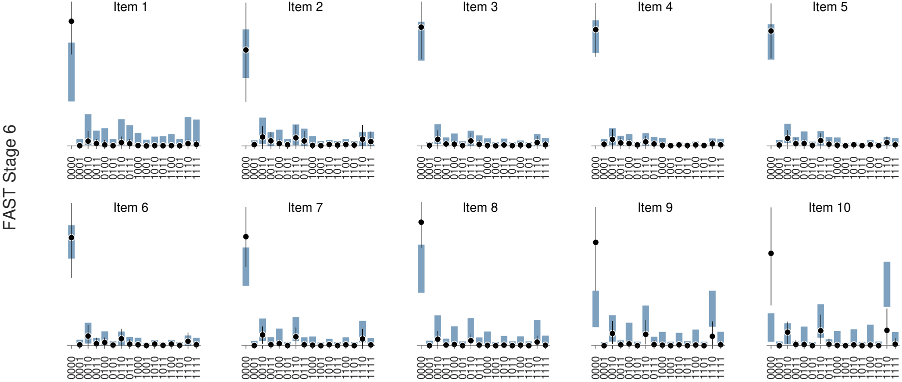
Descriptive Adequacy for Stage 5

- Allowing for each item to have its own parameterization achieves descriptive adequacy



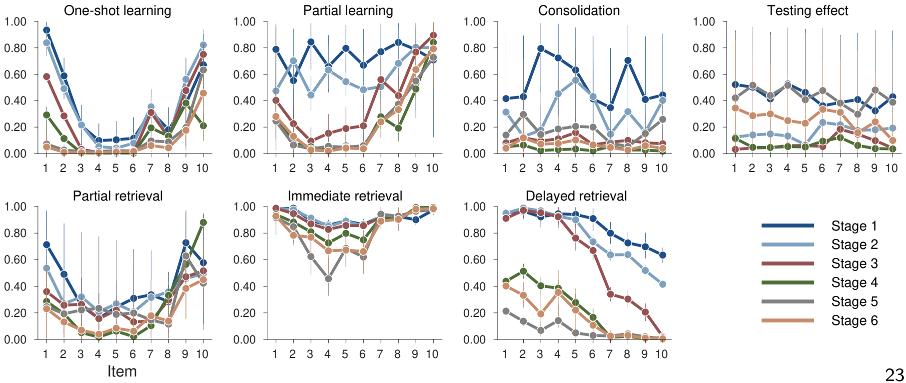
Descriptive Adequacy for Stage 6

- Allowing for each item to have its own parameterization achieves descriptive adequacy



Parameter Inferences

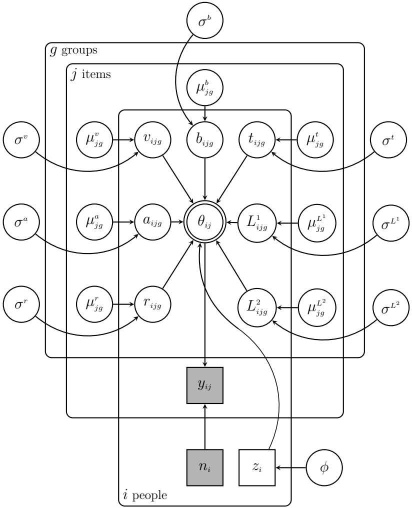
- Although the parameters for each item position are inferred independently, they show clear theoretically-interpretable regularities
  - Serial position effects for immediate retrieval (t, L_1), and decaying primacy effects for delayed retrieval (L_2)
  - Possible serial position effects for learning (a, r, v), except for constant testing effects (b)



Independent Item
Latent-Mixture Model

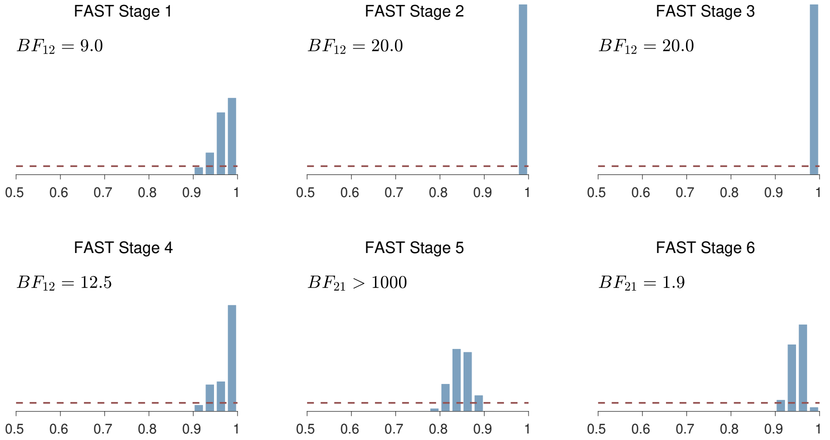
Latent Mixture Model

Allow for two different subgroups, with each person assigned to one, and a base-rate of $\phi \sim \mathrm{Uniform}(0.5, 1)$ for the majority group



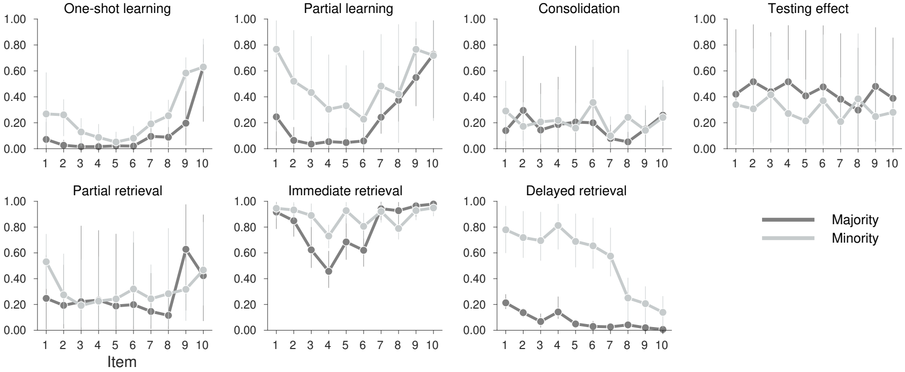
Evidence for Subgroups

- There is evidence that FAST stages 1–4 have only one group, but stage 5 has subgroups, and stage 6 may have subgroups



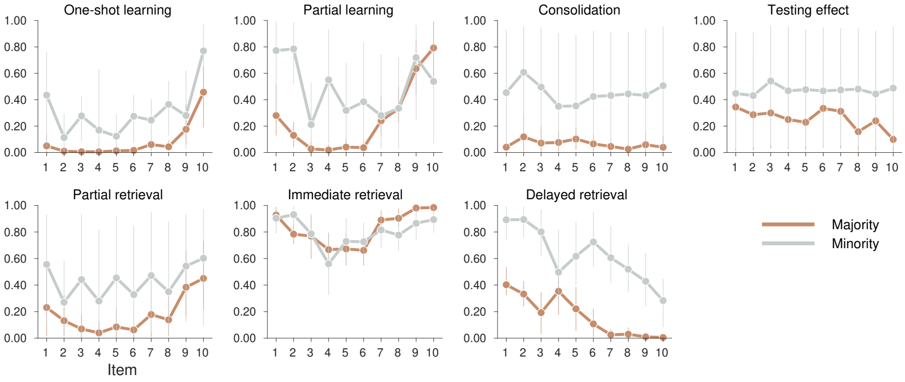
Subgroups in Stage 5

- The minority subgroup, with about 15% of the patients, performs much better than the others



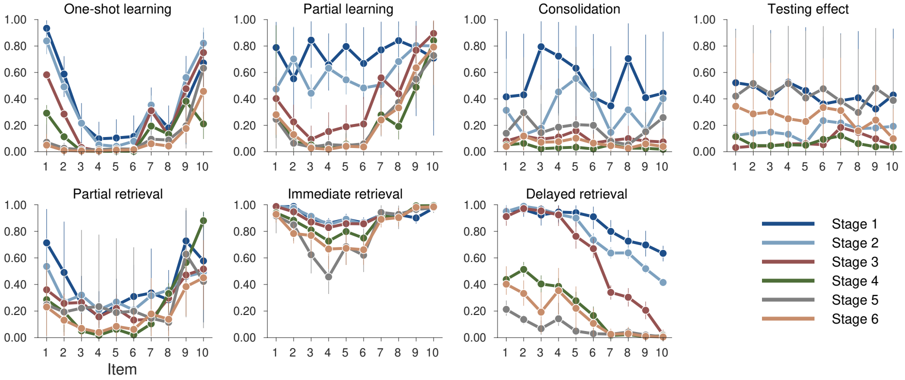
Subgroups in Stage 6

- The minority subgroup, with about 5% of the patients, performs much better than the others



Parameter Inferences

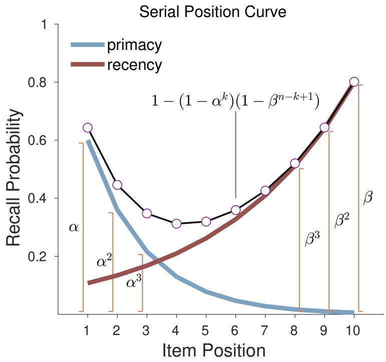
- Serial position effects for immediate retrieval (t, L_1), and decaying primacy effects for delayed retrieval (L_2)
- Possible serial position effects for learning (a, r, v), except for constant testing effects (b)

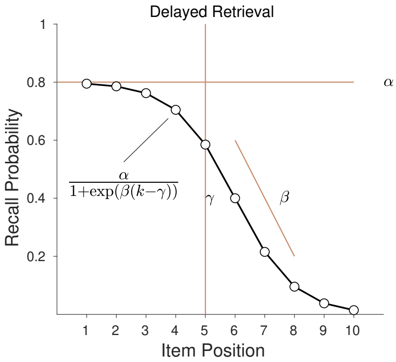


Hierarchical Item Model

Theoretical Extensions to Batchelder Model

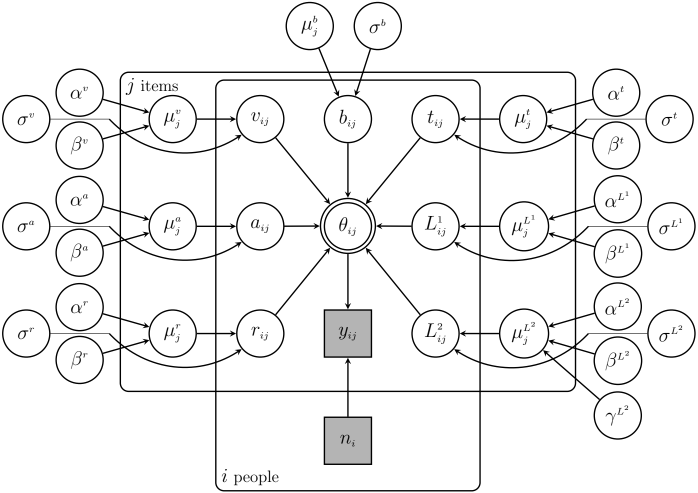
- Hierarchical model of item parameters in terms of their positions
  - Serial position curve model for encoding parameters a, t, r, and immediate retrieval parameters t, L_1
  - Logistic model of delayed retrieval parameter L_2
  - Constant testing effect learning parameter b





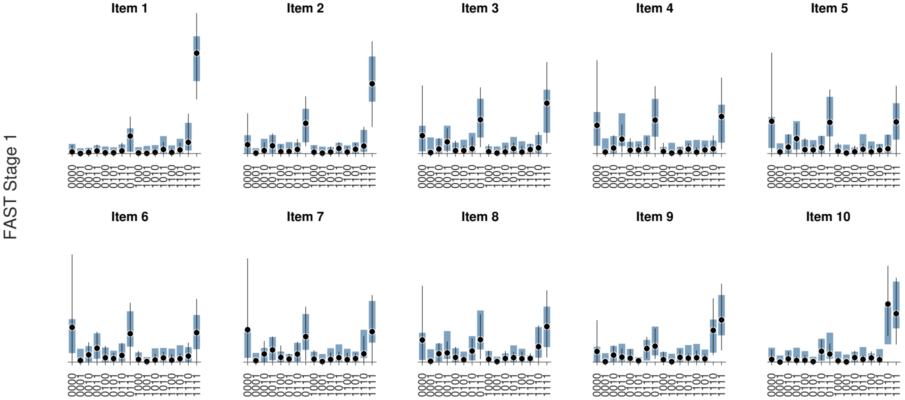
Hierarchical Item Model

Assume individual differences from a truncated Gaussian for each parameter, and now allow independent parameters for each item position



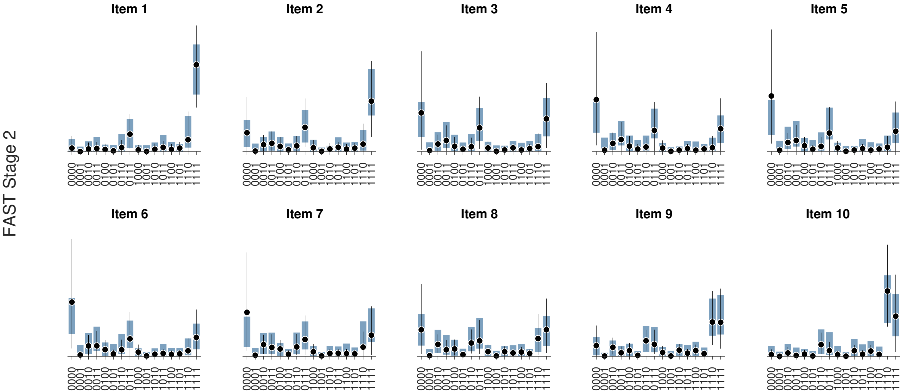
Descriptive Adequacy for Stage 1

- The theoretically-extended model maintains descriptive adequacy



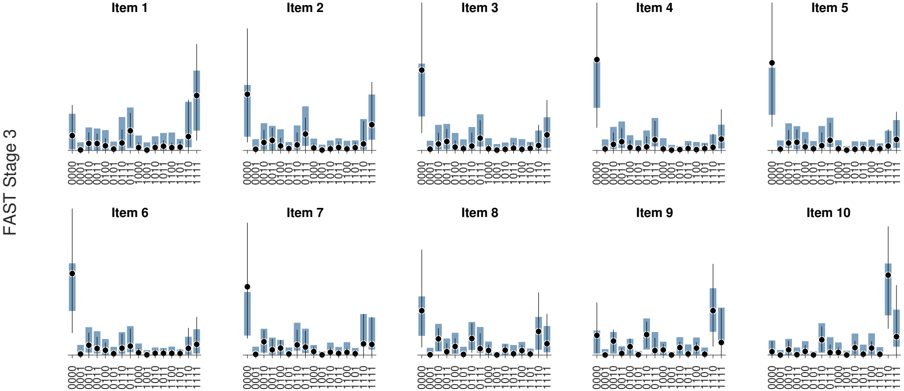
Descriptive Adequacy for Stage 2

- The theoretically-extended model maintains descriptive adequacy



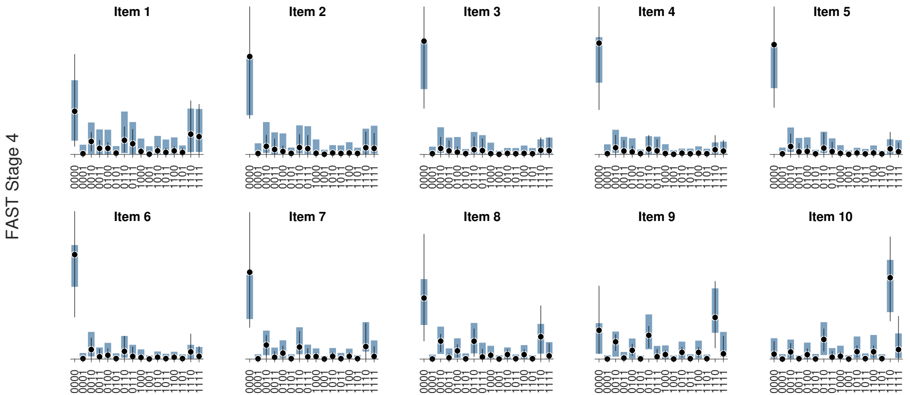
Descriptive Adequacy for Stage 3

- The theoretically-extended model maintains descriptive adequacy



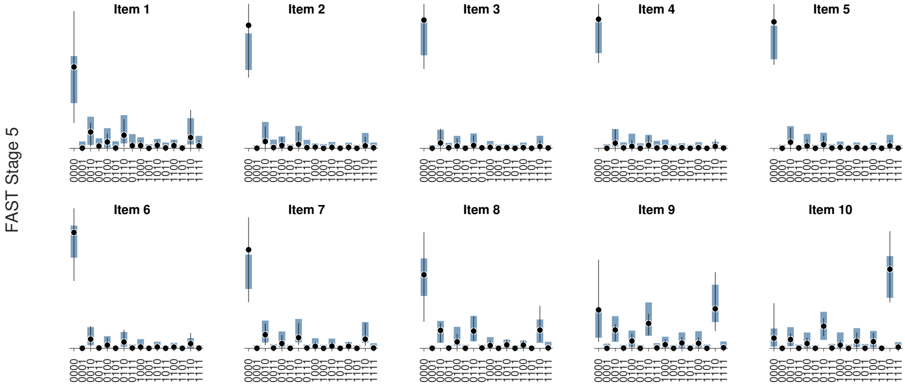
Descriptive Adequacy for Stage 4

- The theoretically-extended model maintains descriptive adequacy



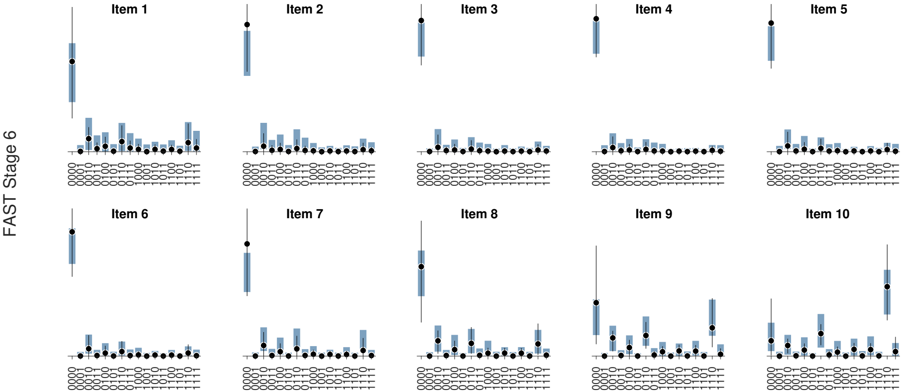
Descriptive Adequacy for Stage 5

- The theoretically-extended model maintains descriptive adequacy



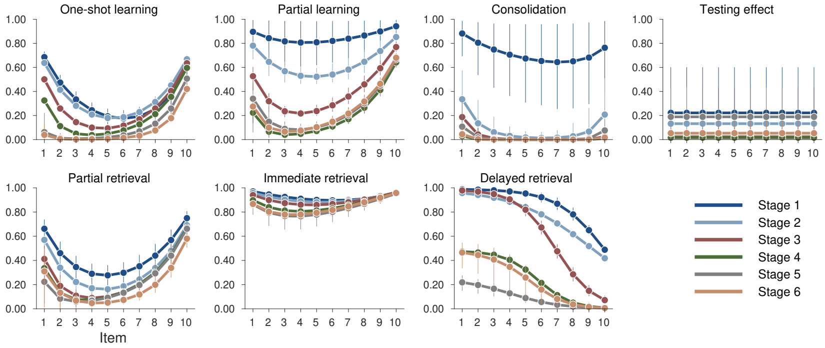
Descriptive Adequacy for Stage 6

- The theoretically-extended model maintains descriptive adequacy



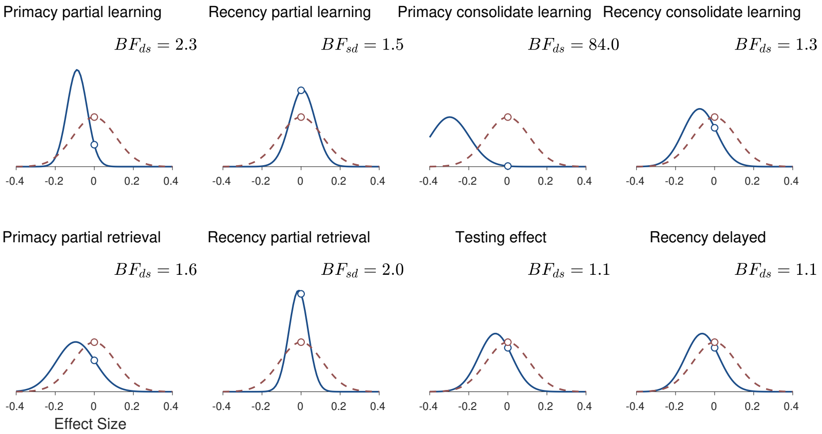
Inferences of Theoretically-Extended Model

- Comparing FAST stage 1 to stage 2 examines the subjective change from cognitively normal to cognitively normal but with a subjective sense of memory impairment
  - No difference in overall recall accuracy, nor in everyday function, for people in these stages



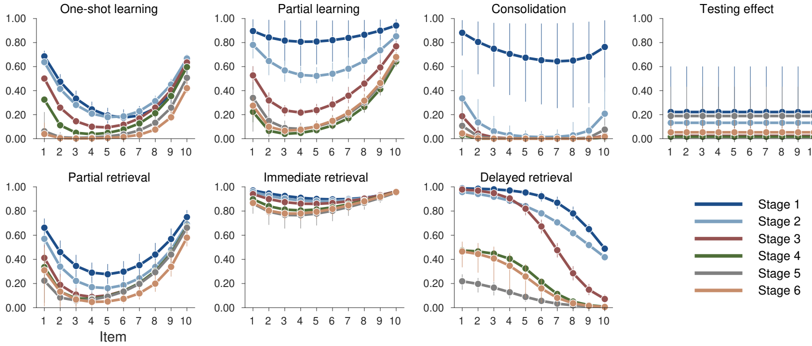
FAST Stage 1 vs 2

- The change in effect size at $\delta = 0$ from prior to posterior gives the Bayes factor for sameness or difference between FAST stage 1 and 2
  - Stage 2 has worse consolidation of partially-learned words at the beginning of the list



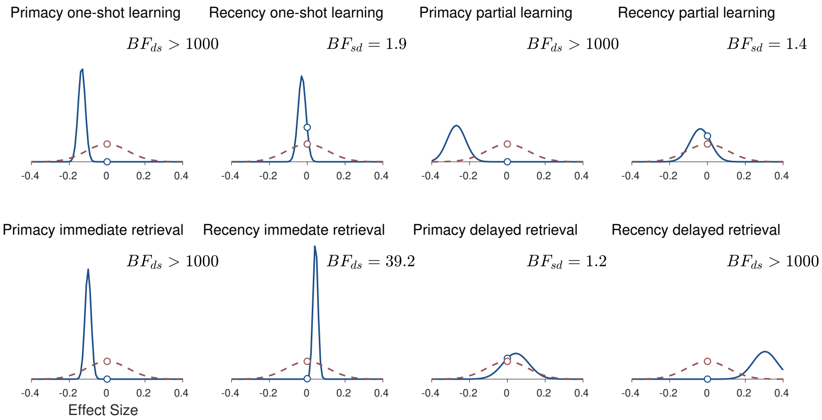
Inferences of Theoretically-Extended Model

- Comparing FAST stage 2 to stage 3 examines the objective change from cognitively normal to cognitively impaired



FAST Stage 2 vs 3

- Learning words presented at the beginning of the list is much worse in stage 3, as is the immediate and delayed recall of later words



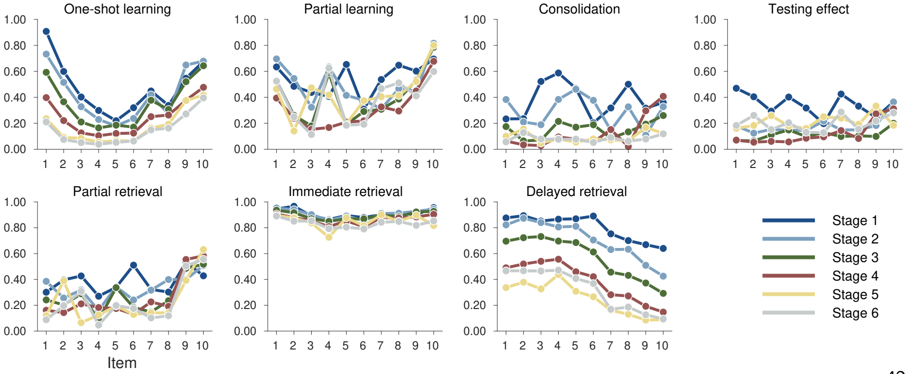
Some Preliminary Conclusions About Memory

- Subjective decline from FAST stage 1 to FAST stage 2 involves difficulties with partial learning
  - Deficits in consolidating encoding of partially-learned words
- More severe objective cognitive impairment to FAST stage 3 and beyond involves deterioriation in long-term memory and rehearsal processes
  - Failure to recall words presented at the beginning of lists in immediate free recall
  - Failure to recall words presented at the end of lists in delayed free recall

Two Final Things

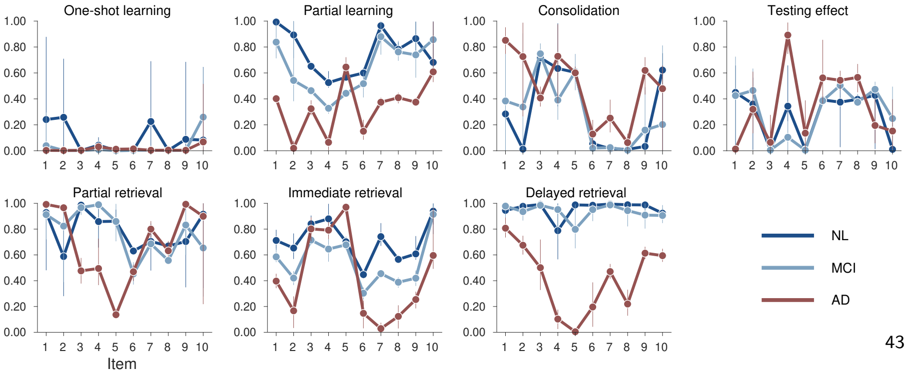
Need to Incorporate Individual Differences

- The inferences are qualitatively different, and less theoretically sensible, if individual differences are removed
- These results are based on aggregating all assessments in each stage, which is equivalent to assuming there are no individual differences for different people in the same FAST stage



Need to Present Items in Same Order

- The inferences are qualitatively different, and **much** less theoretically sensible, for alternative clinical tests that present words in different orders
- These results are based on ADNI data involving cognitively normal, mildly cognitively impaired, and Alzheimer's disease individuals tested using the ADAS-Cog test



Conclusion

Generative Models and Bayesian Methods

- Case study is an example of the benefits of **generative models** of cognition and the use of **Bayesian methods** of inference (Lee, 2018)
- Generative probabilistic models of cognition
  - Force assumptions to be part of the model, saying how psychological parameters and processes generate data
  - Make models theoretically richer, and force the complete quantification of their predictions
- Bayesian methods allow rich and creative cognitive models to be explored
  - Can always, in principle, apply any generative probabilistic model to data to make inferences in the same way
  - Always represent uncertainty about models and parameters, controlling for complexity in the exploration

Thanks!

References

Alexander, G. E., Satalich, T. A., Shankle, W. R., & Batchelder, W. H. (2016). A cognitive psychometric model for the psychodiagnostic assessment of memory-related deficits. *Psychological Assessment*, *28*, 279.

Lee, M. D. (2018). Bayesian methods in cognitive modeling. In J. Wixted & E.-J. Wagenmakers (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience. Volume 5: Methodology* (Fourth ed., pp. 37–84). John Wiley & Sons.