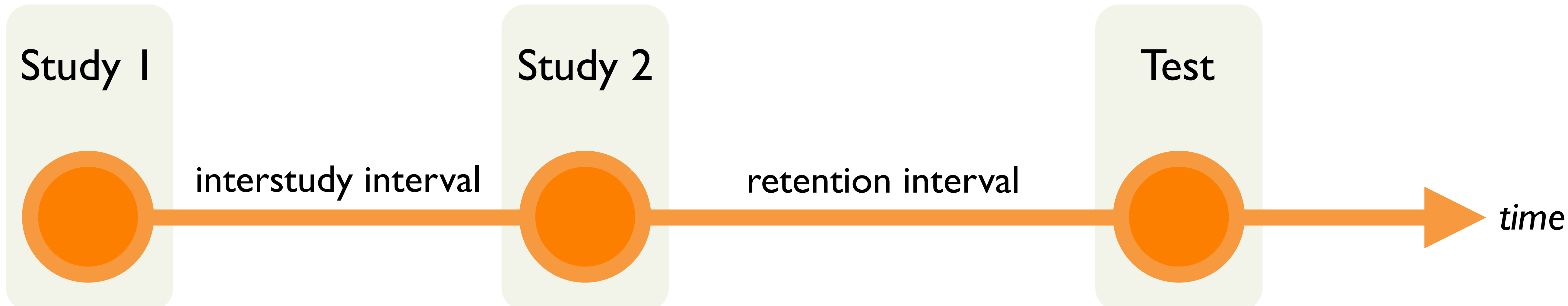
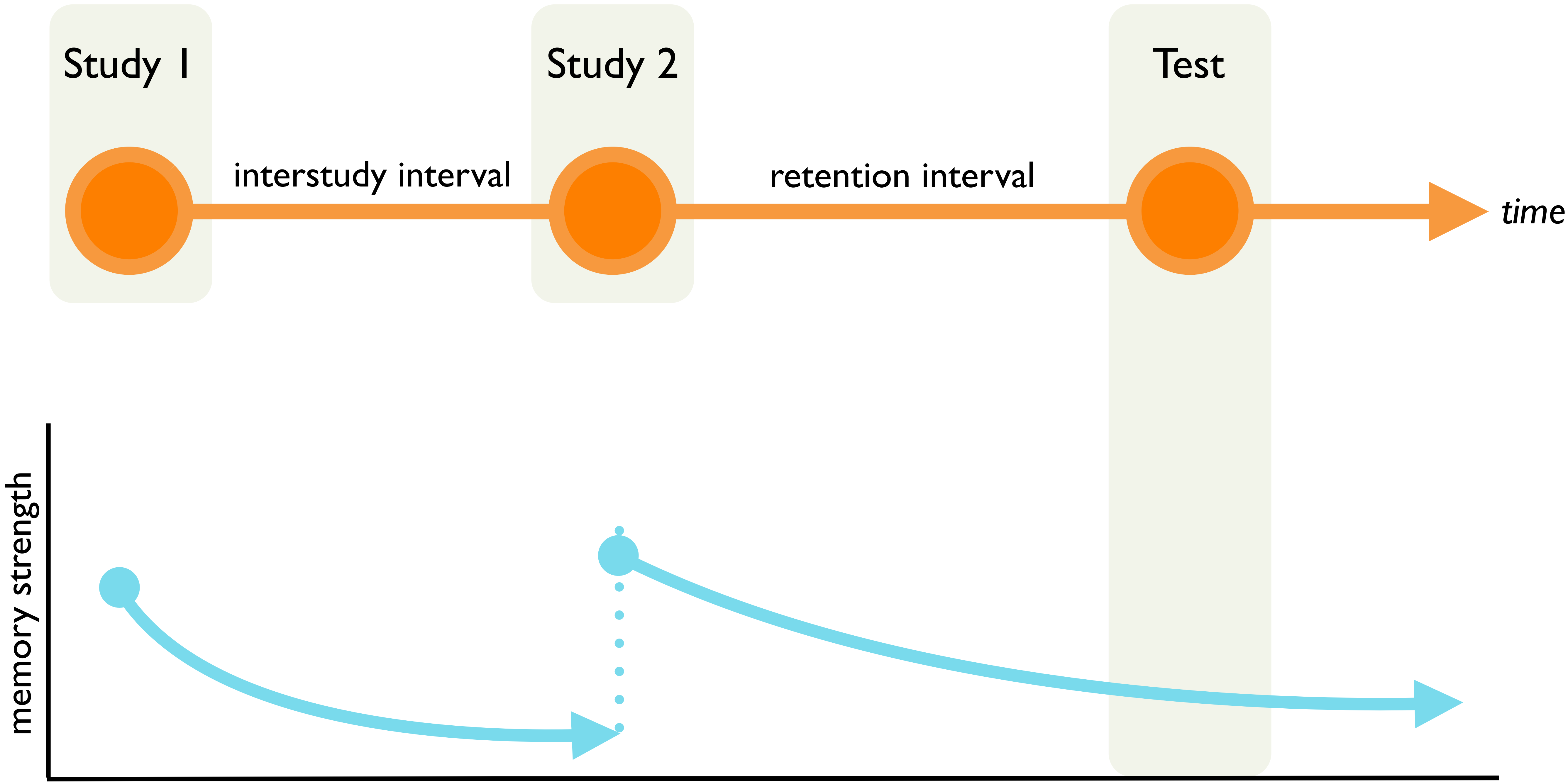
Inferring History-Dependent Memory Strength via Collaborative Filtering

Robert Lindsey, Jeff Shroyer, Michael Mozer
University of Colorado
December 8, 2012

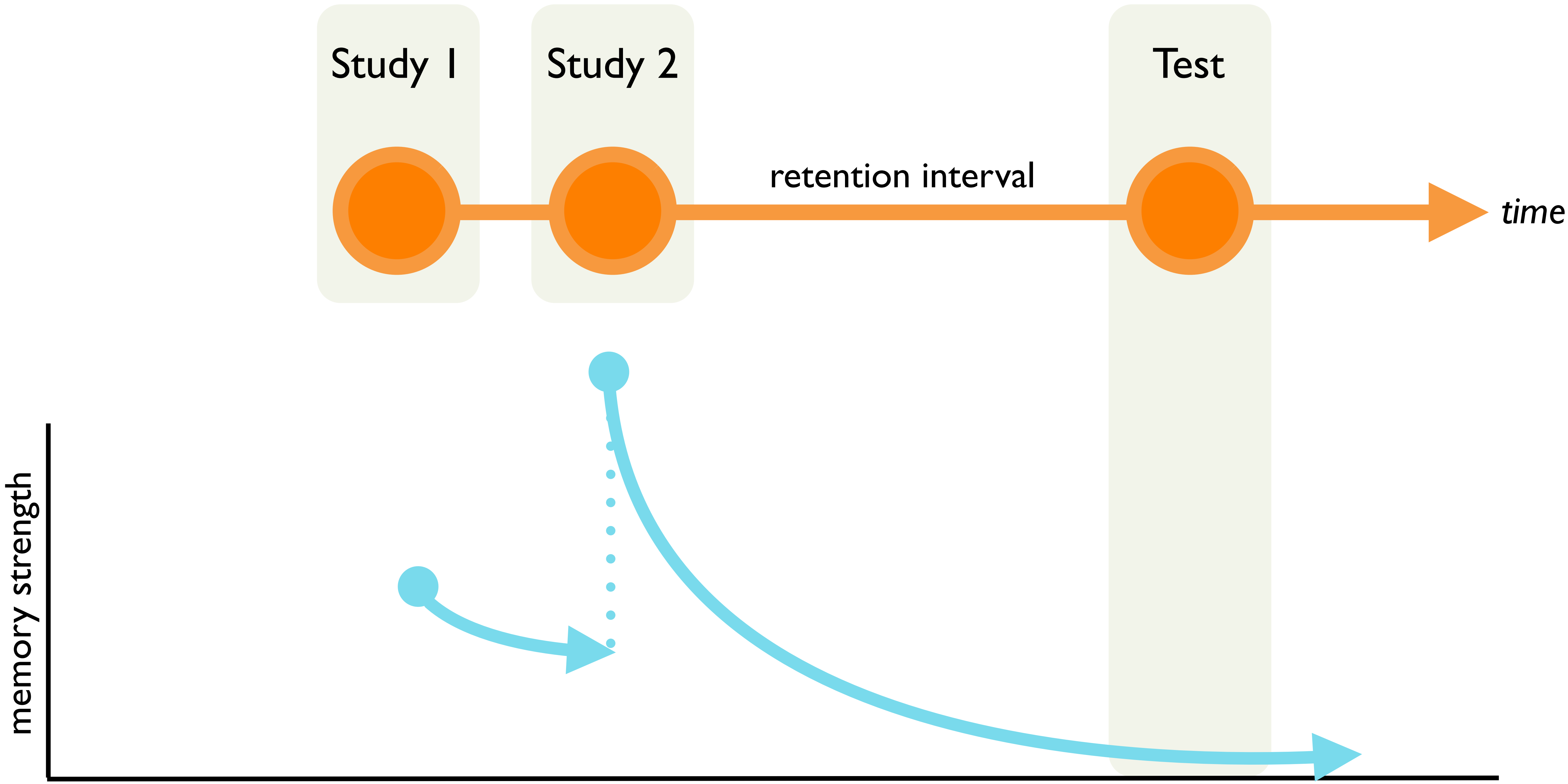
Retrieval probability is strongly influenced by the temporal spacing of study opportunities

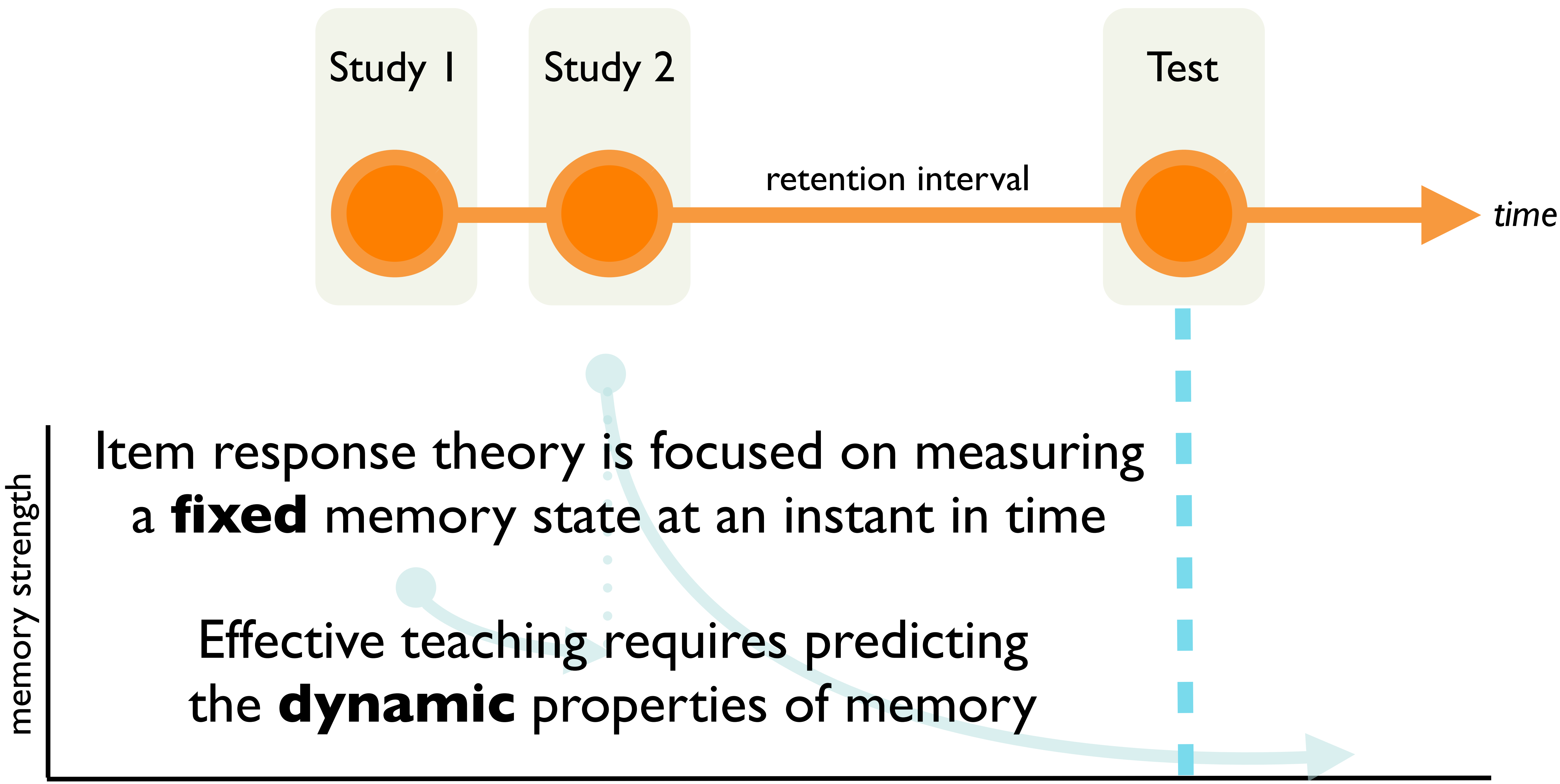


Retrieval probability is strongly influenced by the temporal spacing of study opportunities



Retrieval probability is strongly influenced by the temporal spacing of study opportunities

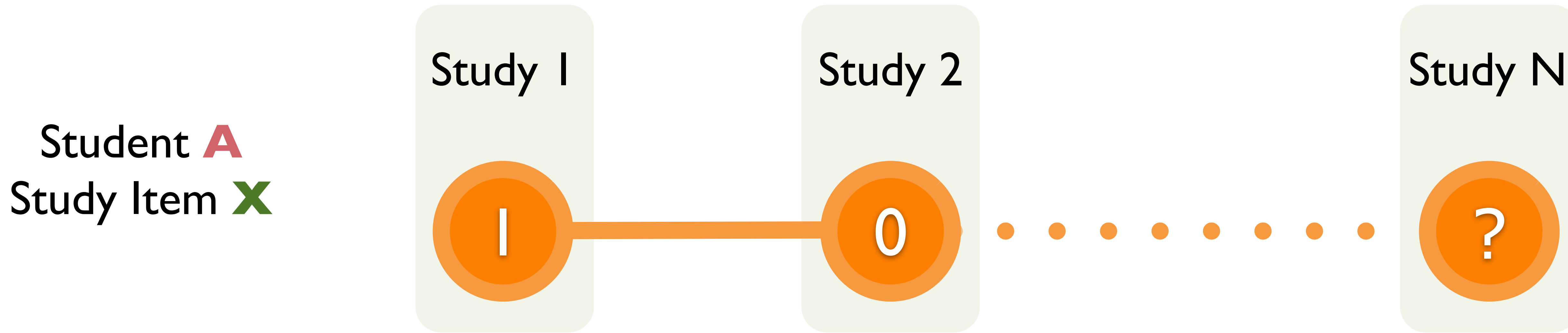




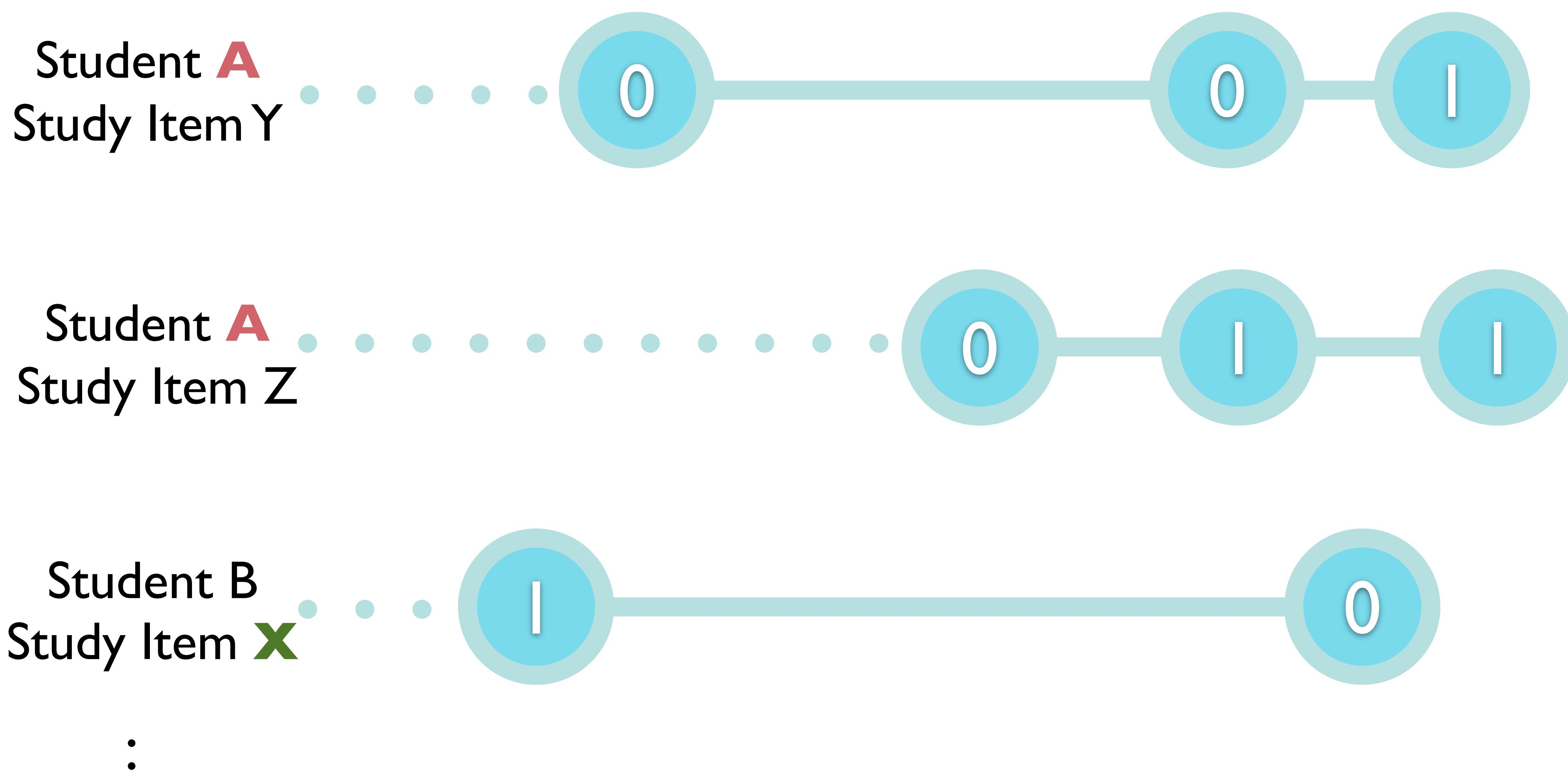
Item response theory is focused on measuring a **fixed** memory state at an instant in time

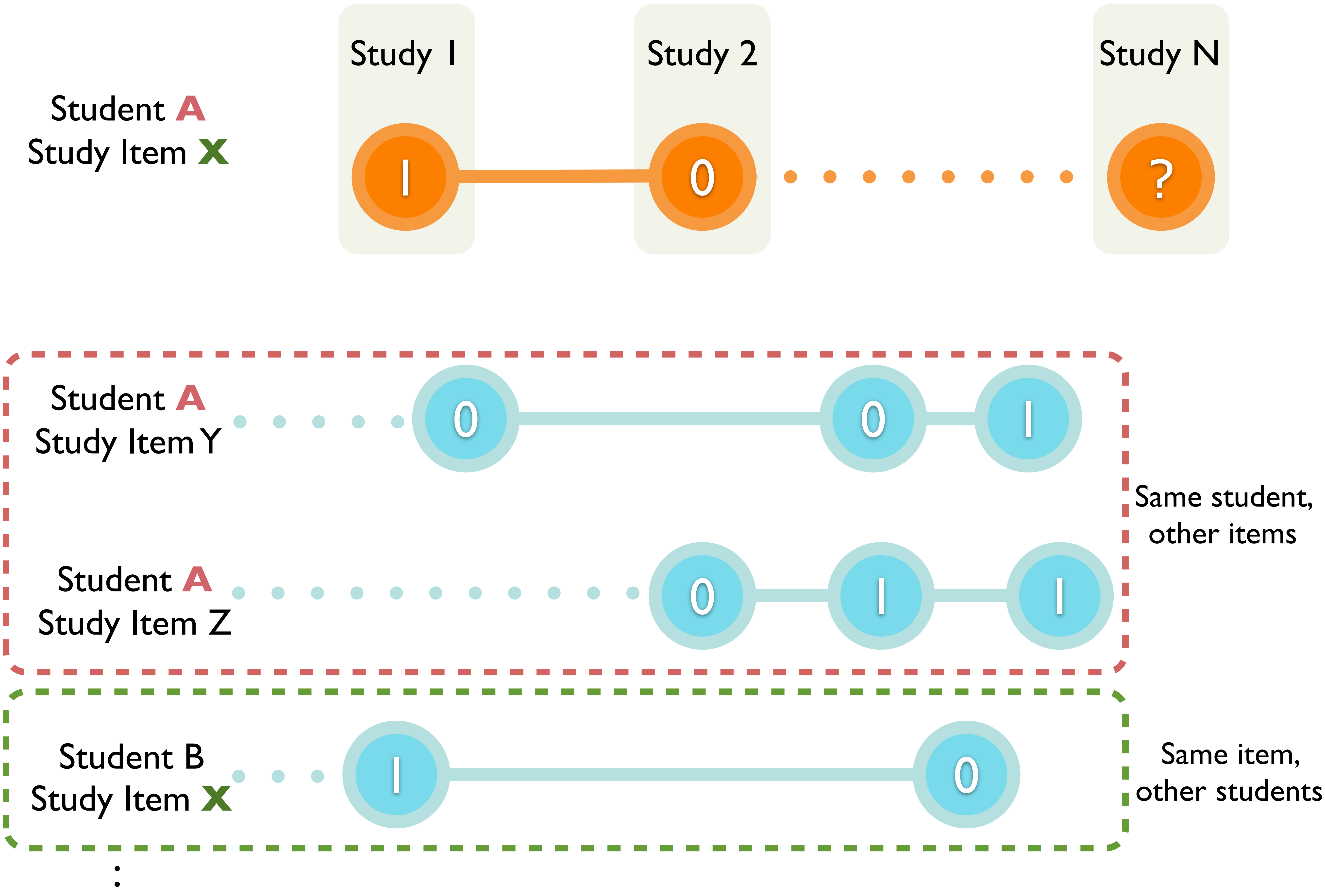
Effective teaching requires predicting the **dynamic** properties of memory

Inference Problem



Training Data




Study 1
Study 2
Study N
Student A
Study Item X
1
0
?
Student A
Study Item Y
0
0
1
Student A
Study Item Z
0
1
1
Same student,
other items
Student B
Study Item X
1
0
Same item,
other students

Naive Approach: Ignore Study History

logistic function

Likelihood

$$\pi_n = \sigma(\alpha_s - \delta_i)$$

ability difficulty

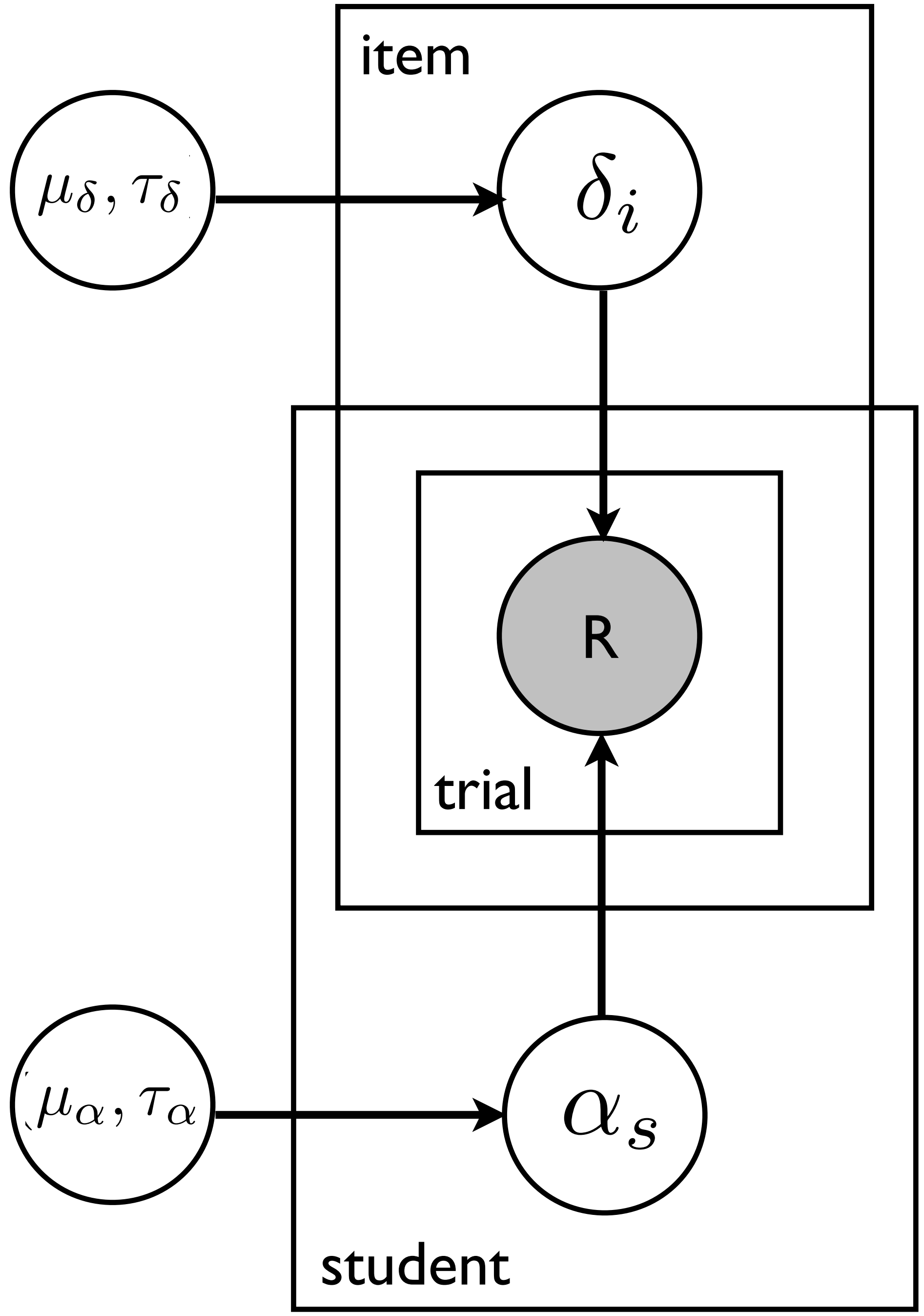
Hierarchical Priors

individuals linked together via shared prior

$$\alpha_s \sim \text{Normal}(\mu_\alpha, \tau_\alpha)$$
$$(\mu_\alpha, \tau_\alpha) \sim \text{NG}(\mu_0^{(\alpha)}, \kappa_0^{(\alpha)}, a_0^{(\alpha)}, b_0^{(\alpha)})$$

$$\delta_i \sim \text{Normal}(\mu_\delta, \tau_\delta)$$
$$(\mu_\delta, \tau_\delta) \sim \text{NG}(\mu_0^{(\delta)}, \kappa_0^{(\delta)}, a_0^{(\delta)}, b_0^{(\delta)})$$

Rasch Model



Additive Factor Model: Binning Approach

Likelihood

logistic function

$$\pi_n = \sigma(\alpha_s - \delta_i + h_{si})$$

$$h_{si} = \sum_k w_k f_{sik}$$

summarization of study history

observed covariate

Hierarchical Priors

individuals linked together via shared prior

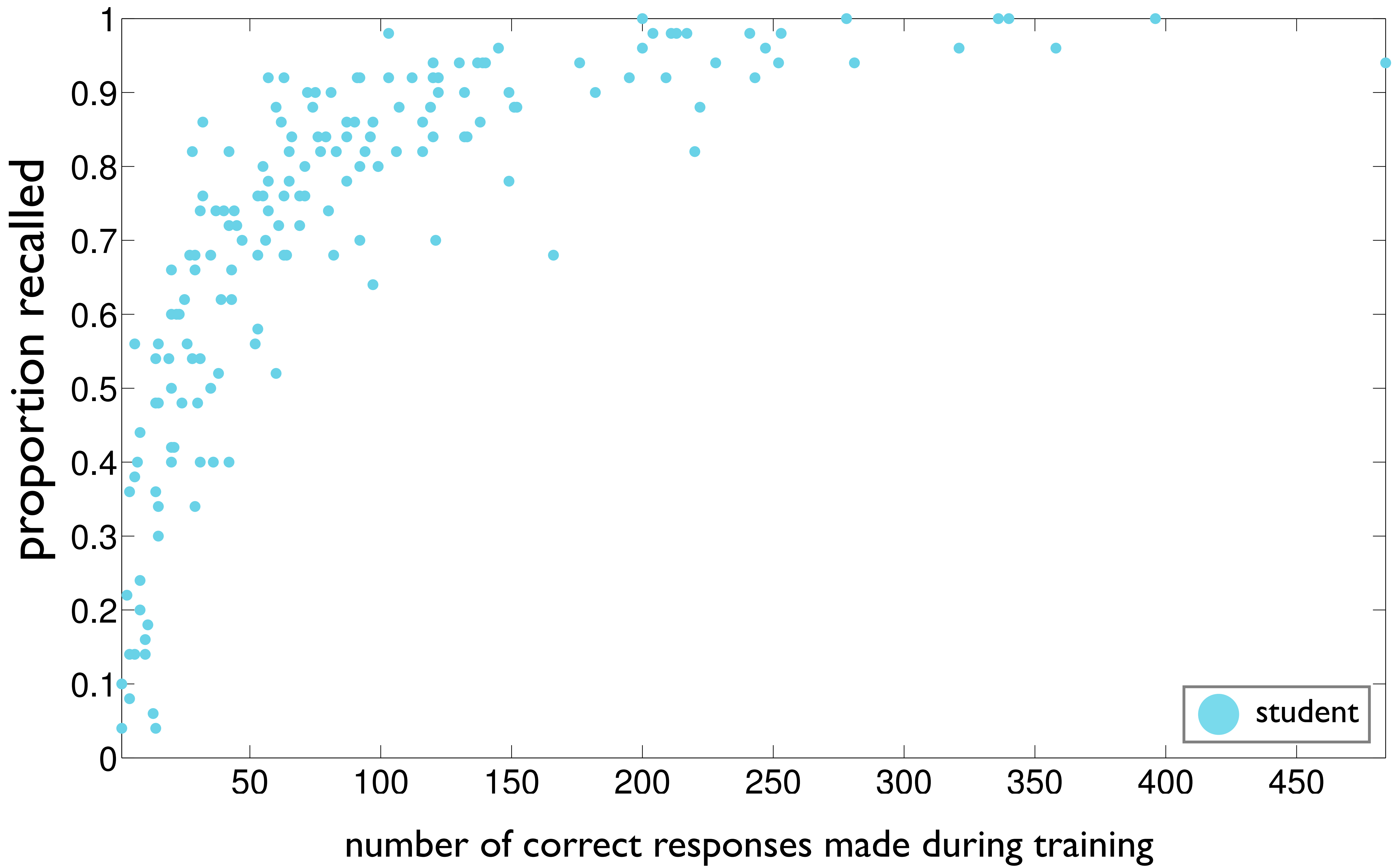
$$\begin{aligned} \alpha_s &\sim \text{Normal}(\mu_\alpha, \tau_\alpha) \\ (\mu_\alpha, \tau_\alpha) &\sim \text{NG}(\mu_0^{(\alpha)}, \kappa_0^{(\alpha)}, a_0^{(\alpha)}, b_0^{(\alpha)}) \\ \\ \delta_i &\sim \text{Normal}(\mu_\delta, \tau_\delta) \\ (\mu_\delta, \tau_\delta) &\sim \text{NG}(\mu_0^{(\delta)}, \kappa_0^{(\delta)}, a_0^{(\delta)}, b_0^{(\delta)}) \end{aligned}$$

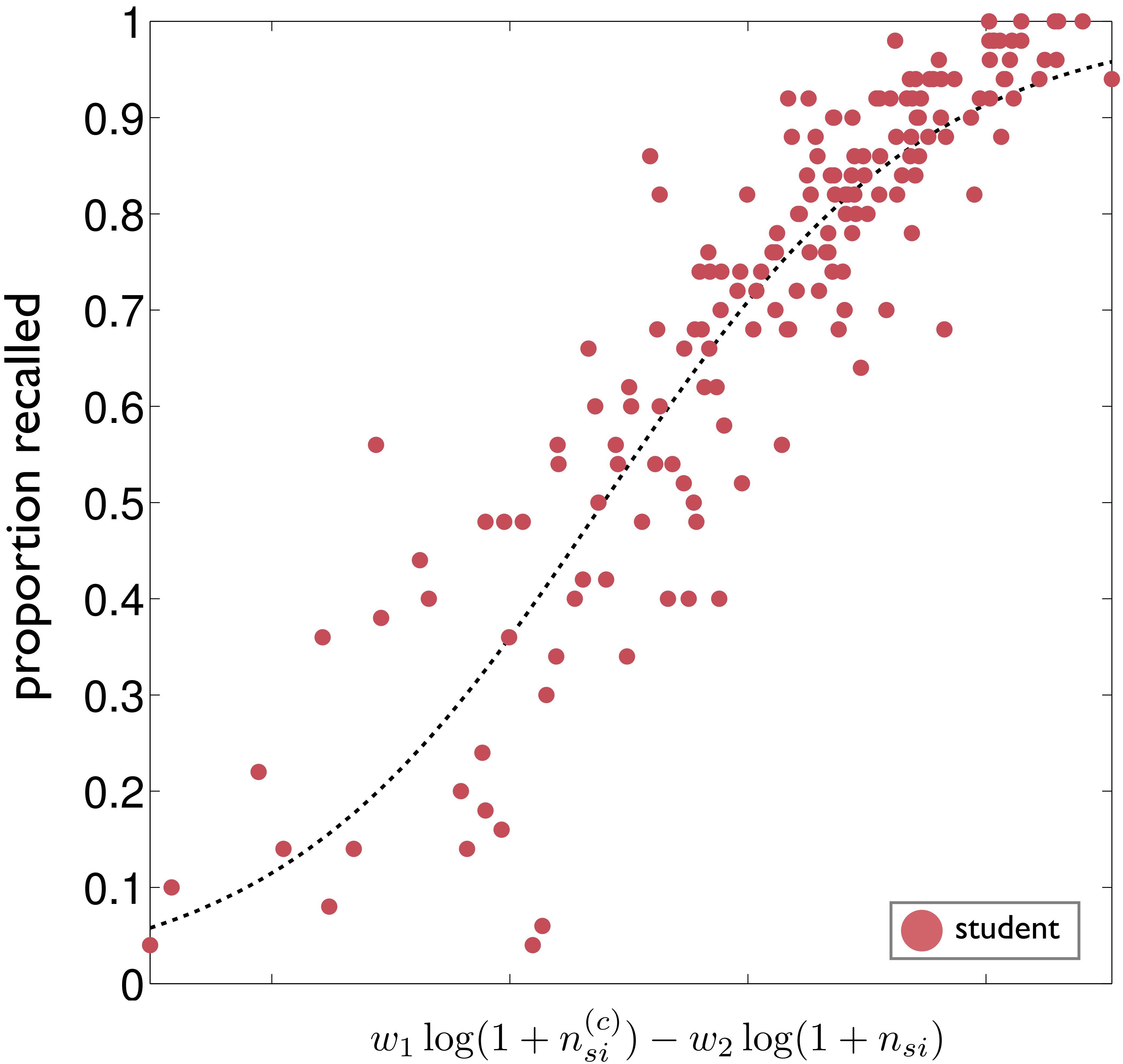
$$f_{sik} = ?$$

$f_{sik} = ?$

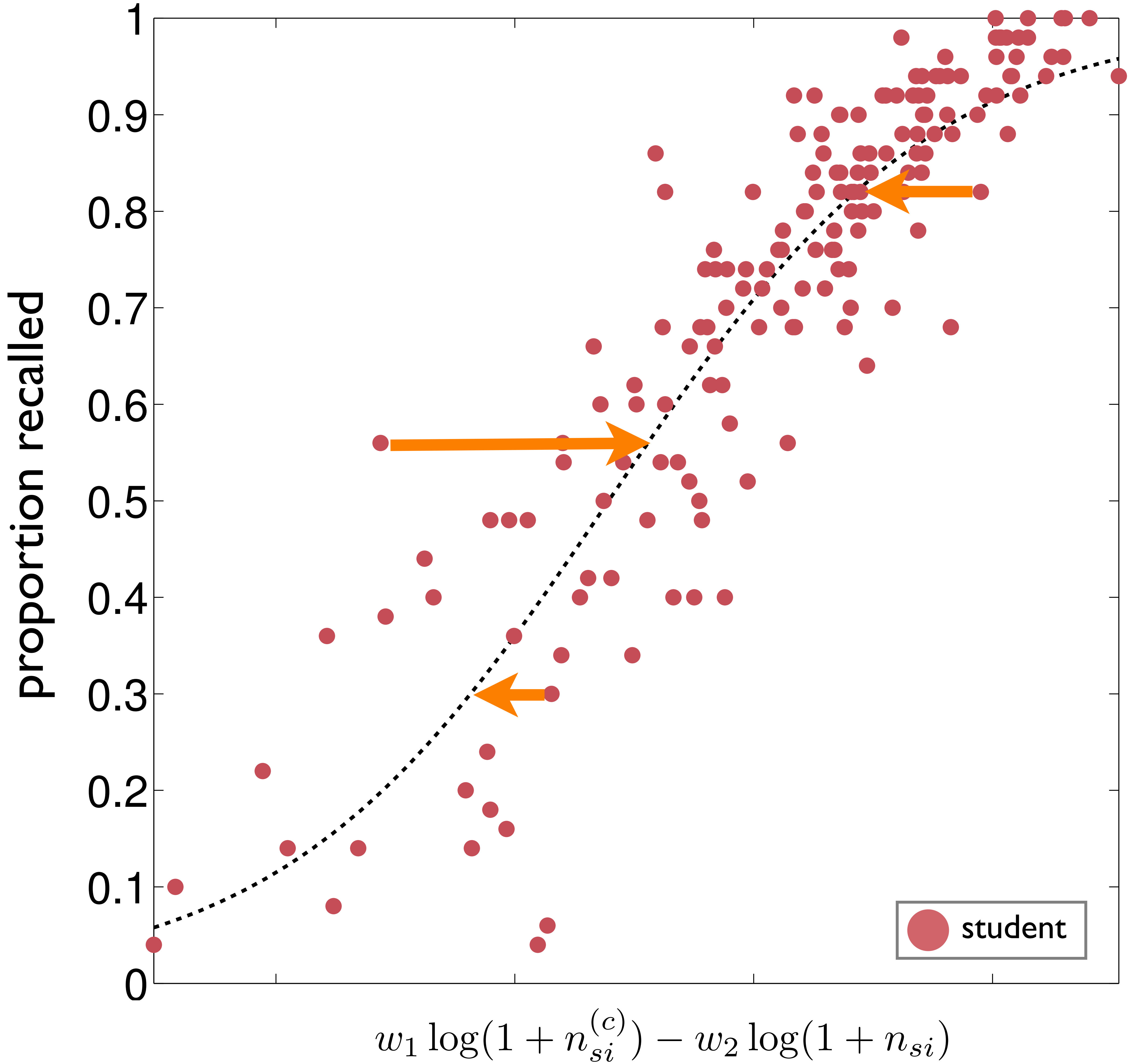
Simple Idea:

count the number of correct responses made and the total number of trials





proportion recalled
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
student
$w_1 \log(1 + n_{si}^{(c)}) - w_2 \log(1 + n_{si})$



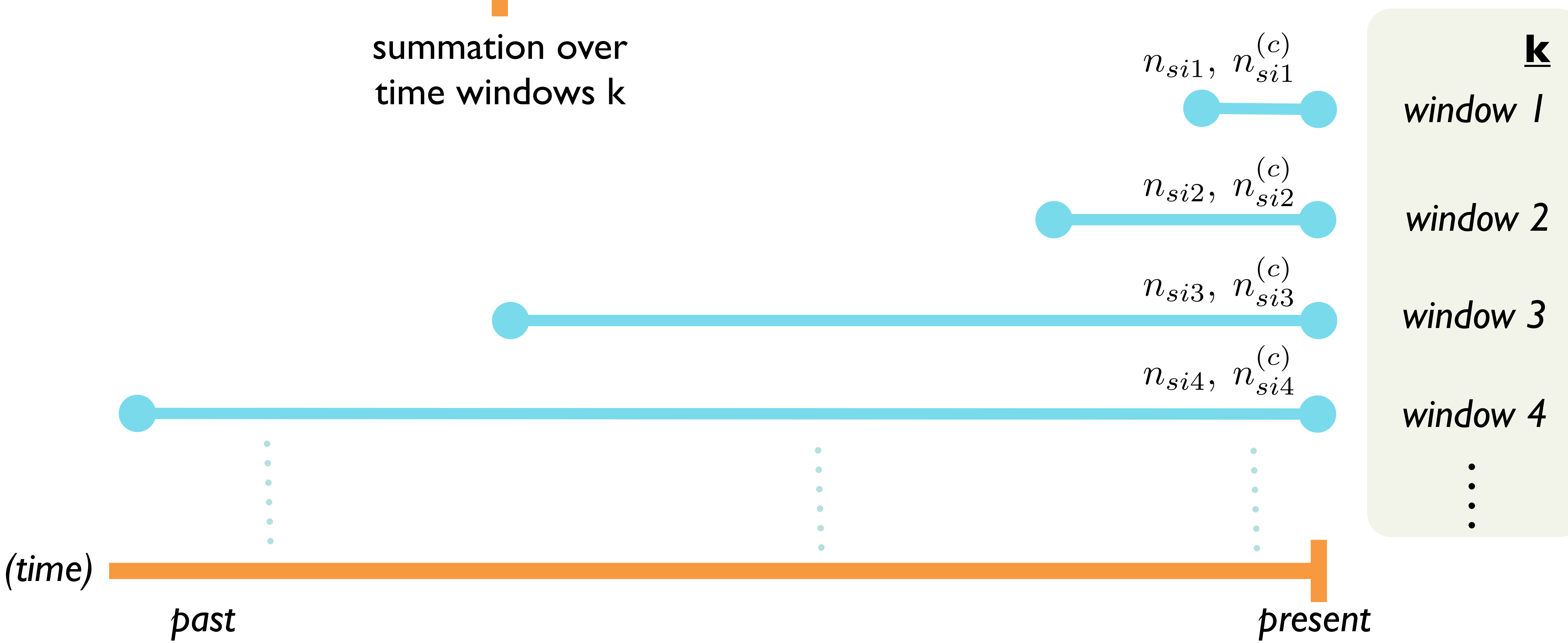
proportion recalled
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
$w_1 \log(1 + n_{si}^{(c)}) - w_2 \log(1 + n_{si})$
student

$$\pi_n = \sigma(\alpha_s - \delta_i + h_{si})$$

counts in a particular time window

$$h_{si} = \sum_k \left[w_{k1}^{(c)} \log(1 + n_{sik}^{(c)}) - w_{k2} \log(1 + n_{sik}) \right]$$

summation over time windows k



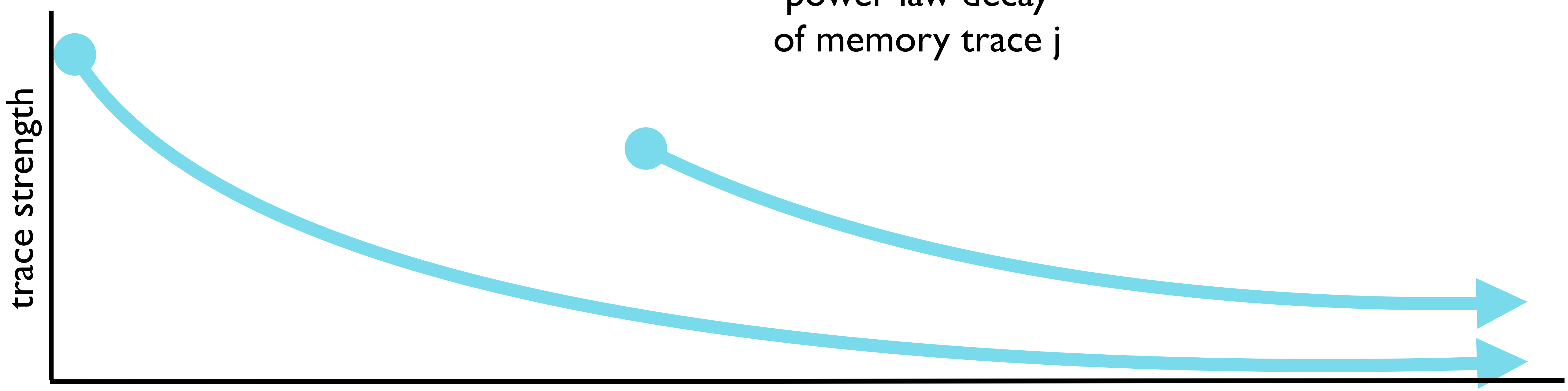
AFM Factor Model: Trace Decay Approach

$$\pi_n = \sigma(\alpha_s - \delta_i + h_{si})$$

summation over
previous
presentations j

$$h_{si} = c \log(1 + \sum_j m_j \, t_{sij}^{-d})$$

power law decay
of memory trace j



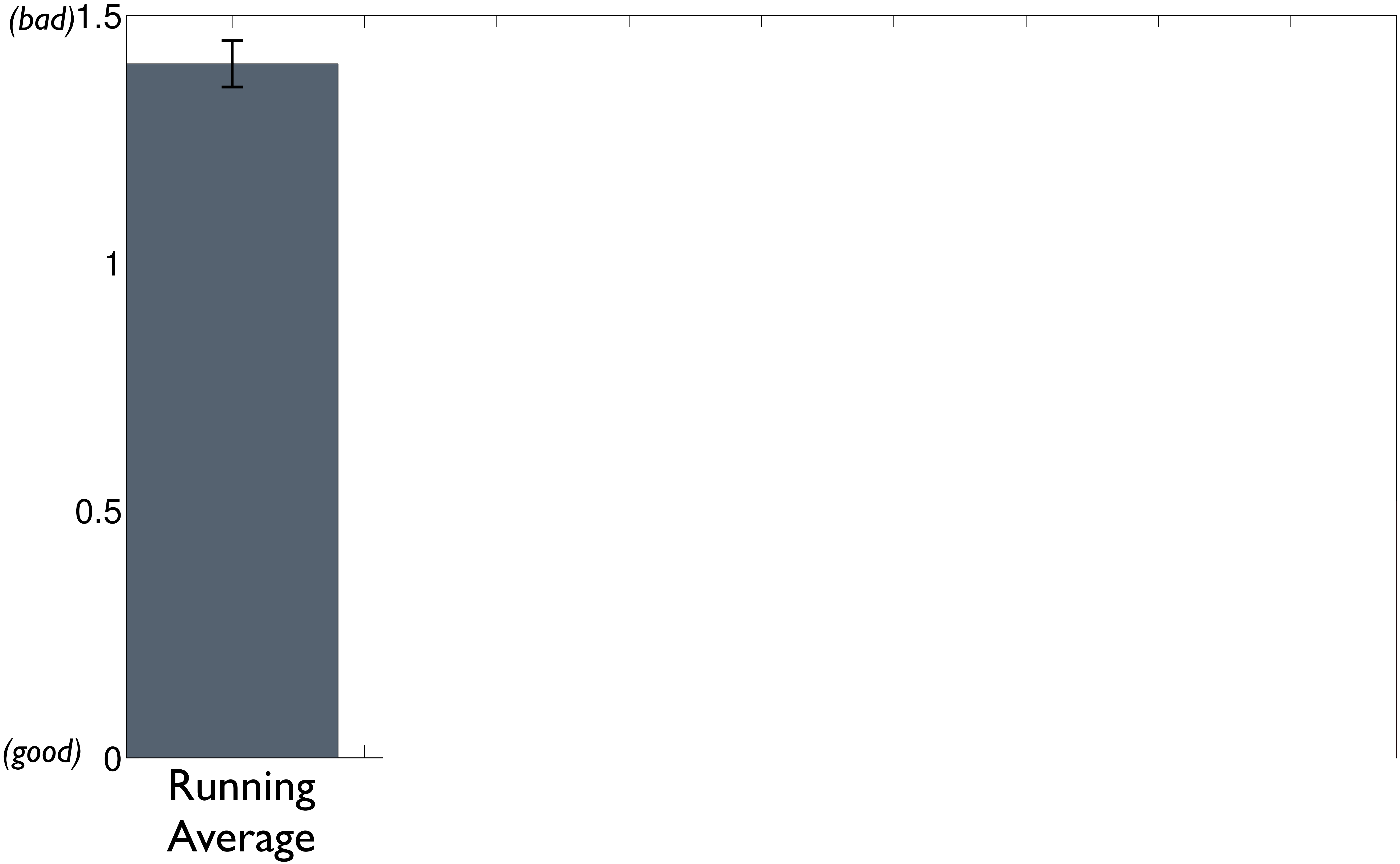
Dataset

- Online vocabulary tutor
- 8th grade Spanish classes
  - 180 students
  - 409 study items
- 14 week study
- 600,000+ study trials

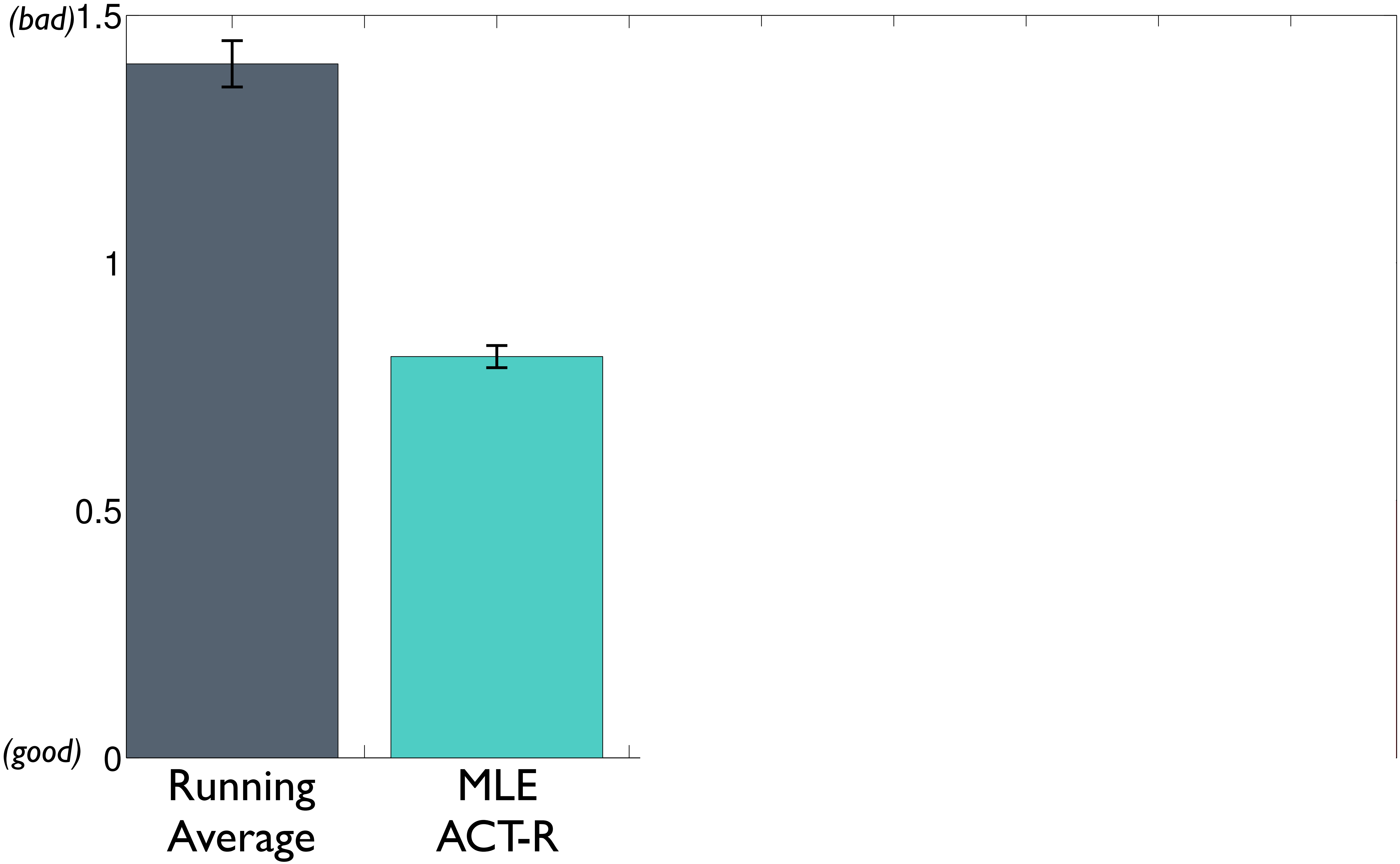
Simulations

Predict recall on day X given data from days 1 through X-1

Average Surprisal

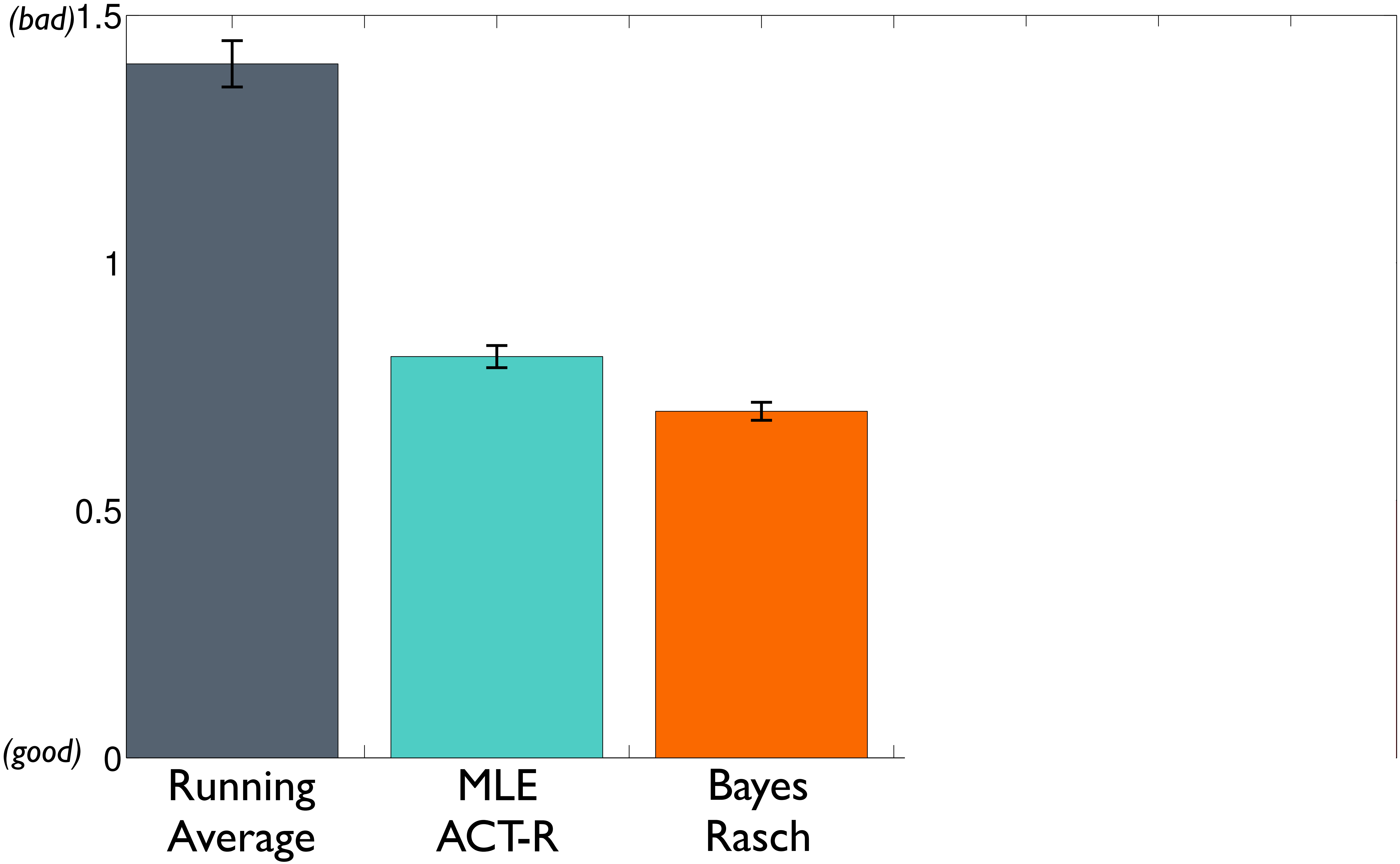


Average Surprisal



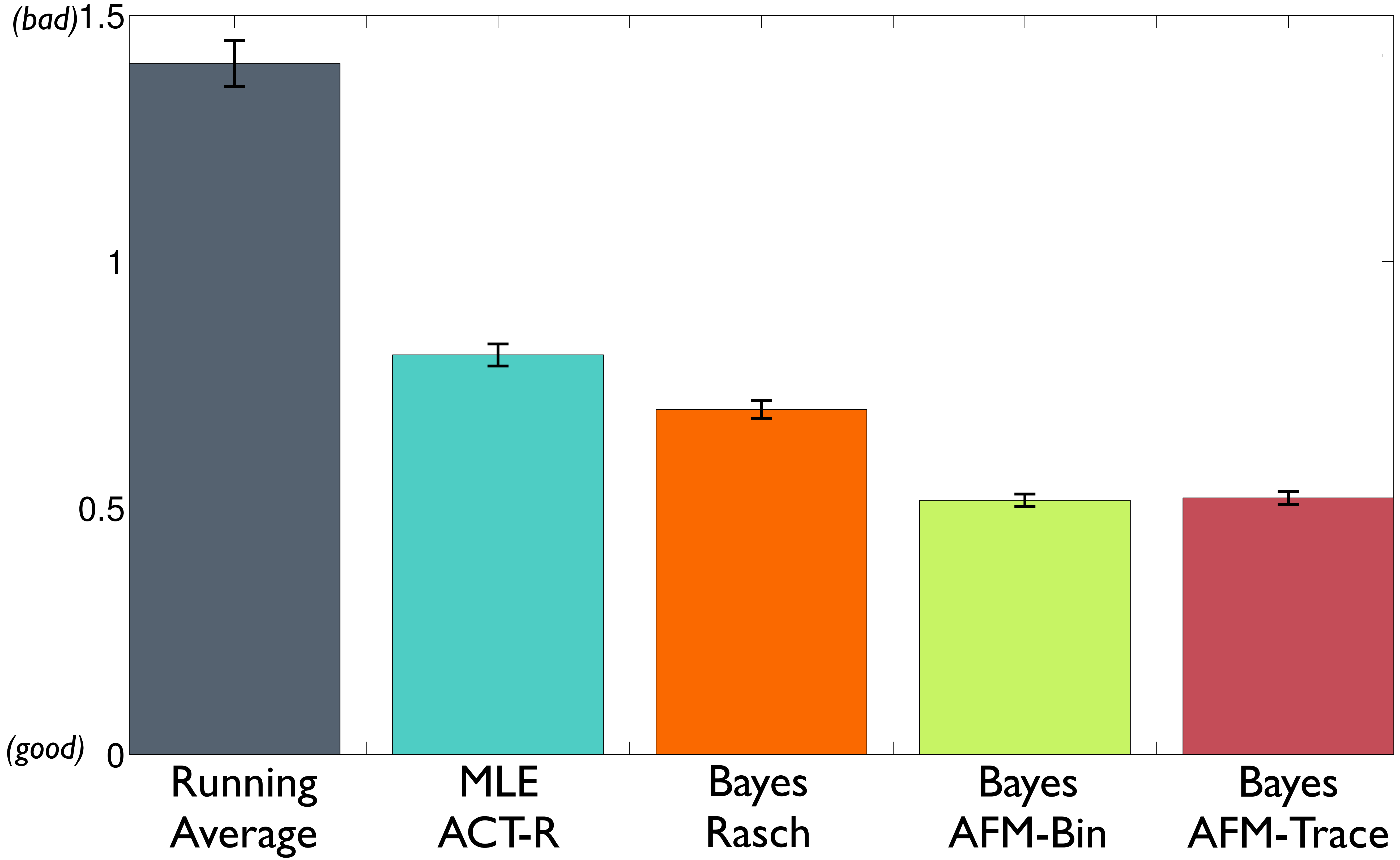
(Pavlik & Anderson, 2008)

Average Surprisal

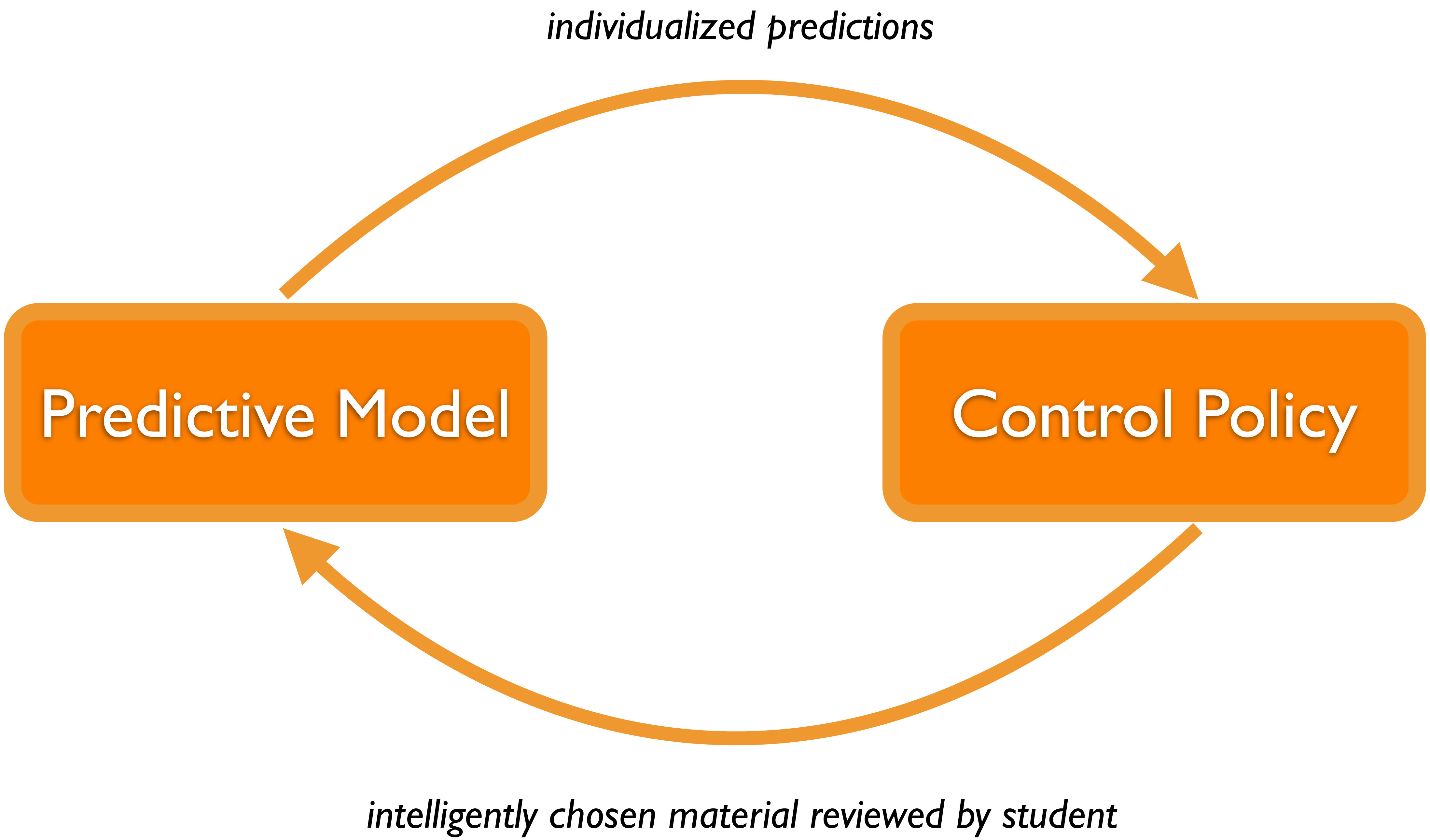


(Pavlik & Anderson, 2008)

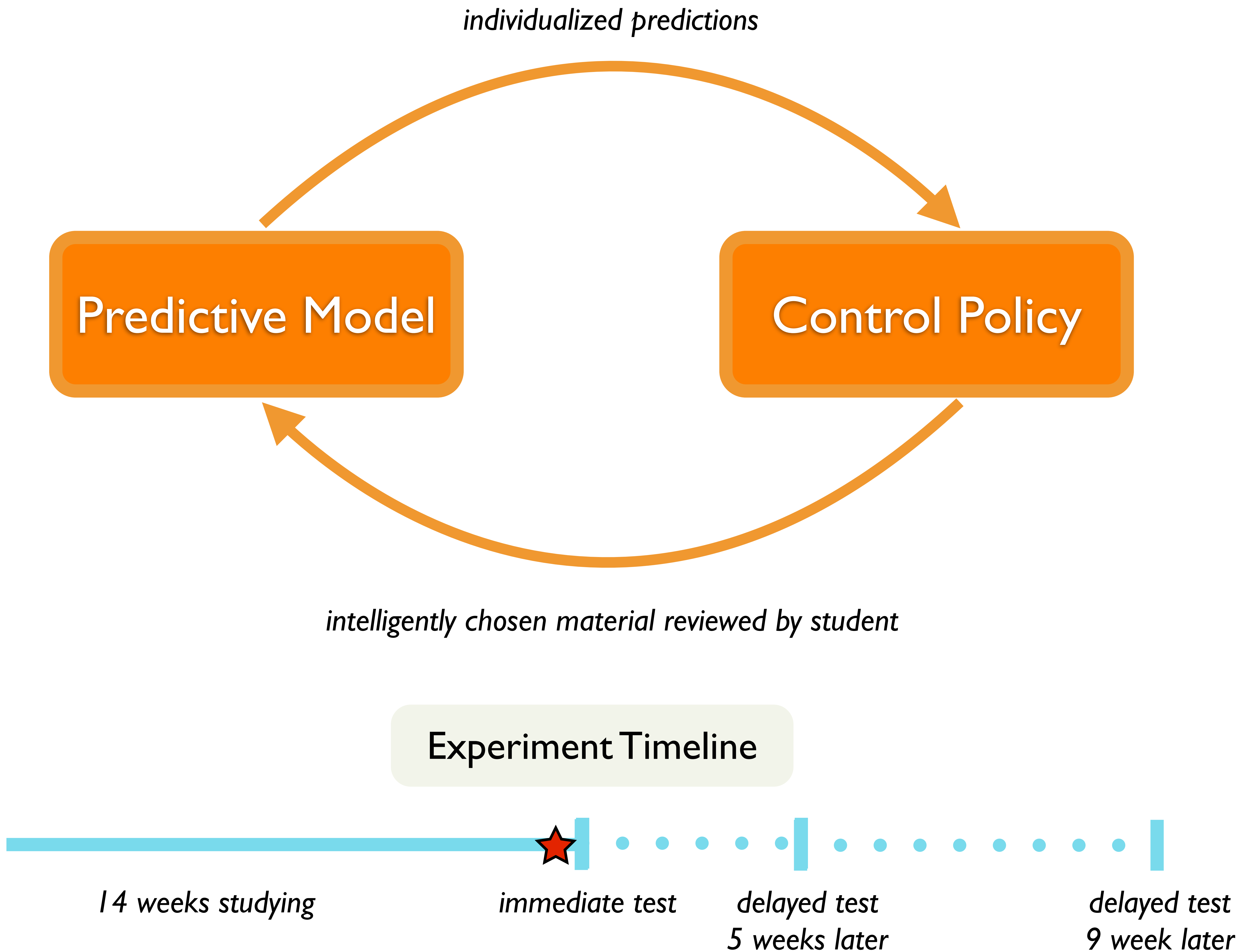
Average Surprisal



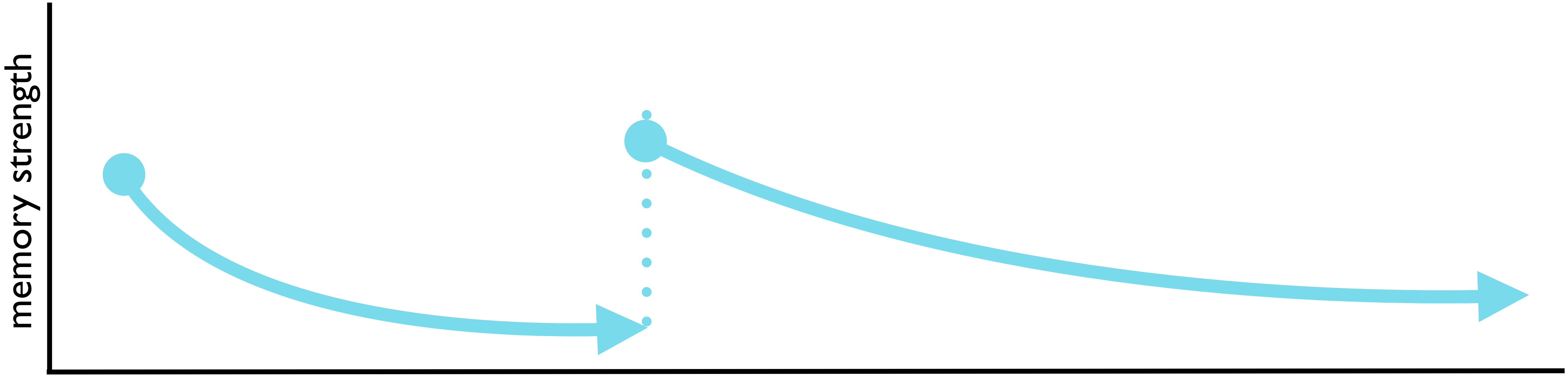
(Pavlik & Anderson, 2008)



To what extent can this scheme improve long term retention?


individualized predictions
Predictive Model
Control Policy
intelligently chosen material reviewed by student
Experiment Timeline
14 weeks studying
immediate test
delayed test
5 weeks later
delayed test
9 week later

Regardless of what's being learned, forgetting occurs.
Accounting for it is important.



Collaborative filtering enables us to make strong inferences despite weak behavioral data

We can use the models to intelligently choose individual items to present to individual students