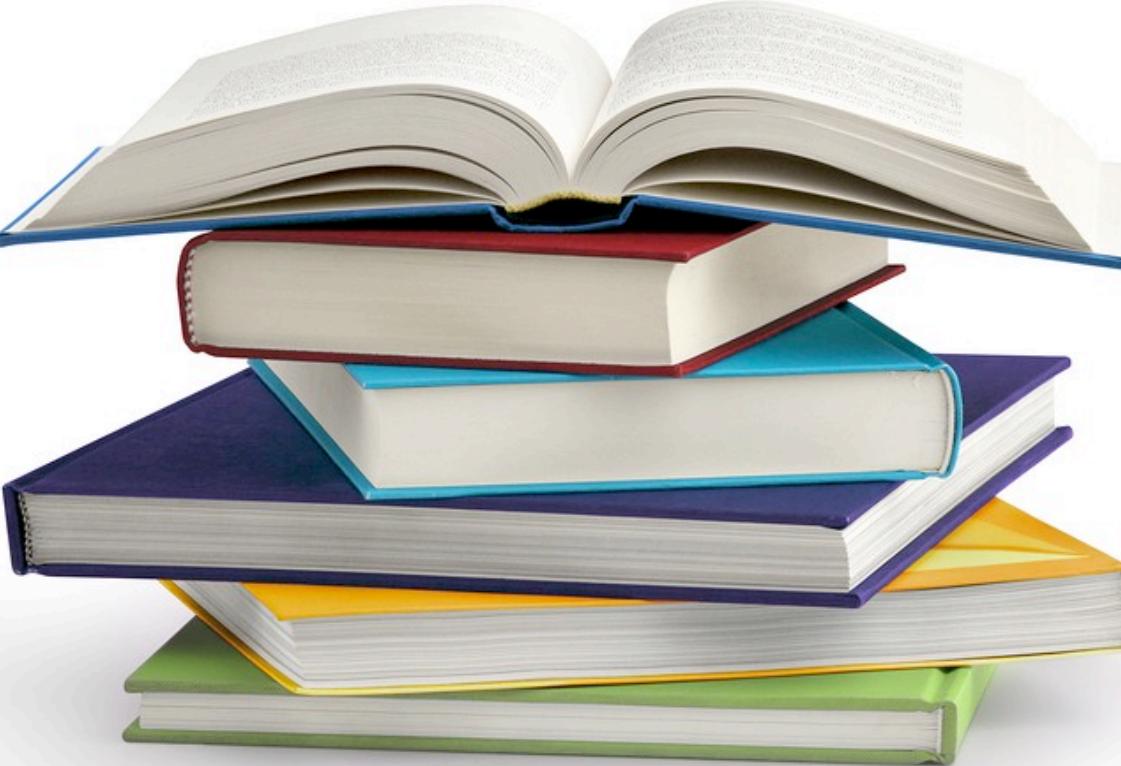


Teaching as hierarchical Bayesian inference

Alicia Chen, Robert Hawkins, Sam Gershman

**CCNL lab meeting
May 25, 2022**

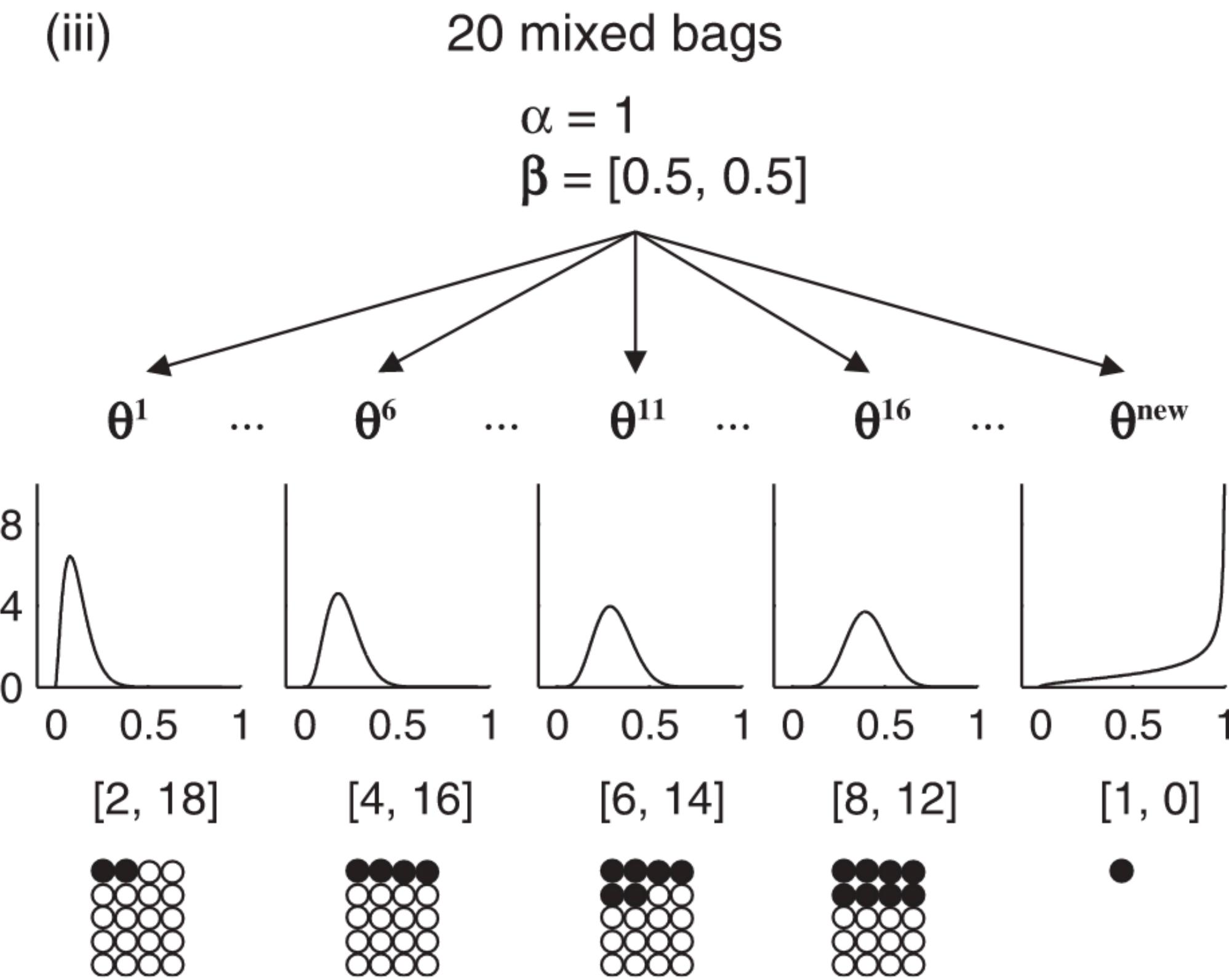
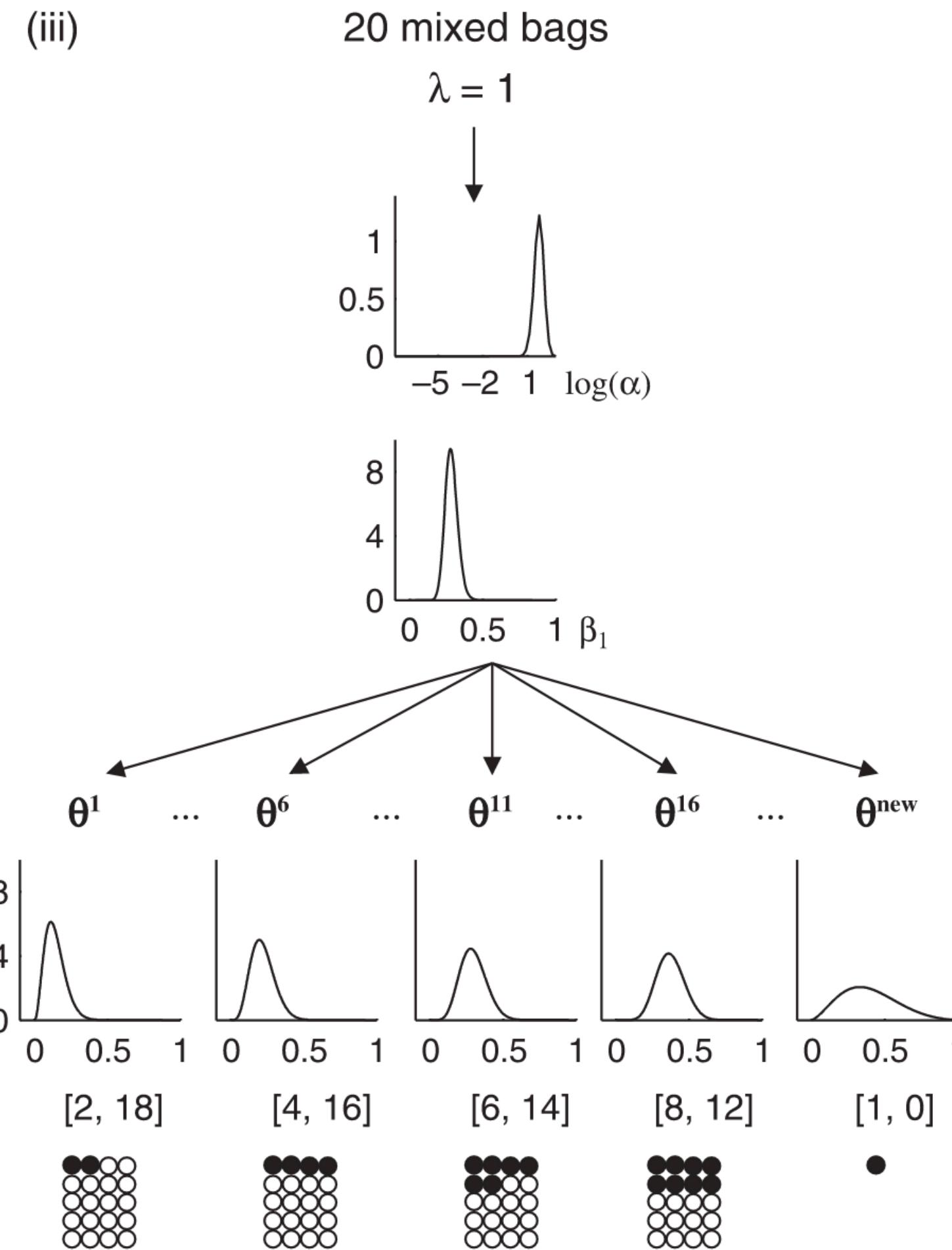


**Question: how do we teach when we
don't know that much about our
students?**

Teaching as learning

- Select data to probe student + learn from student's feedback before teaching (Ho et al. 2017; Fan et al. 2018; Wang, Wu, & Goodman, 2022; Lewandowski et al., 2022)
- Conduct “assessments” by sampling directly from student’s distribution (Liu & Frank, 2018)
- Balance asking questions with teaching; teaching as planning (Popp & Gureckis, 2020)

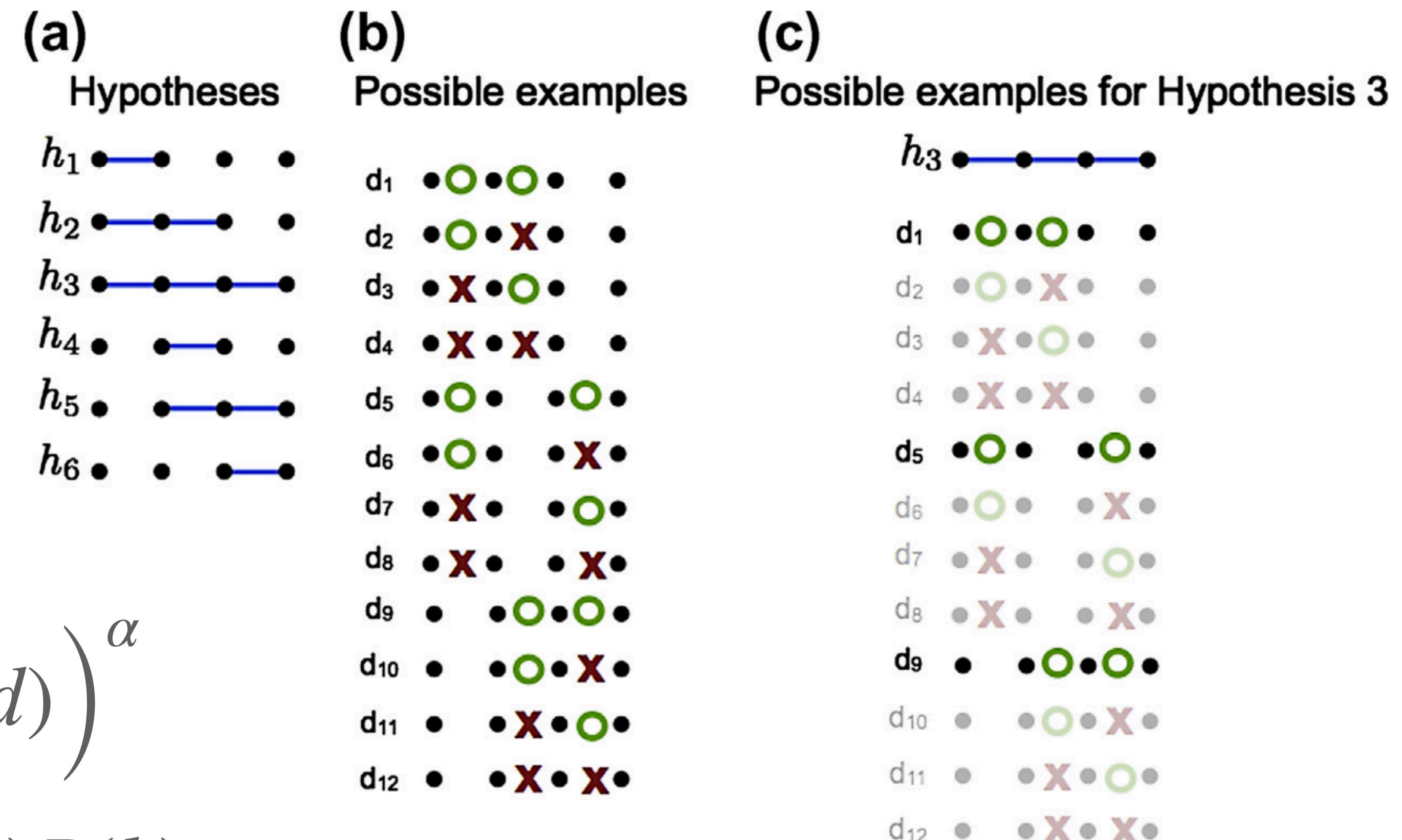
Adding a level of uncertainty



Modeling pedagogical reasoning

$$P_{\text{teacher}}(d|h) \propto \left(P_{\text{learner}}(h|d) \right)^\alpha$$

$$P_{\text{learner}}(h|d) \propto P_{\text{teacher}}(d|h)P(h)$$

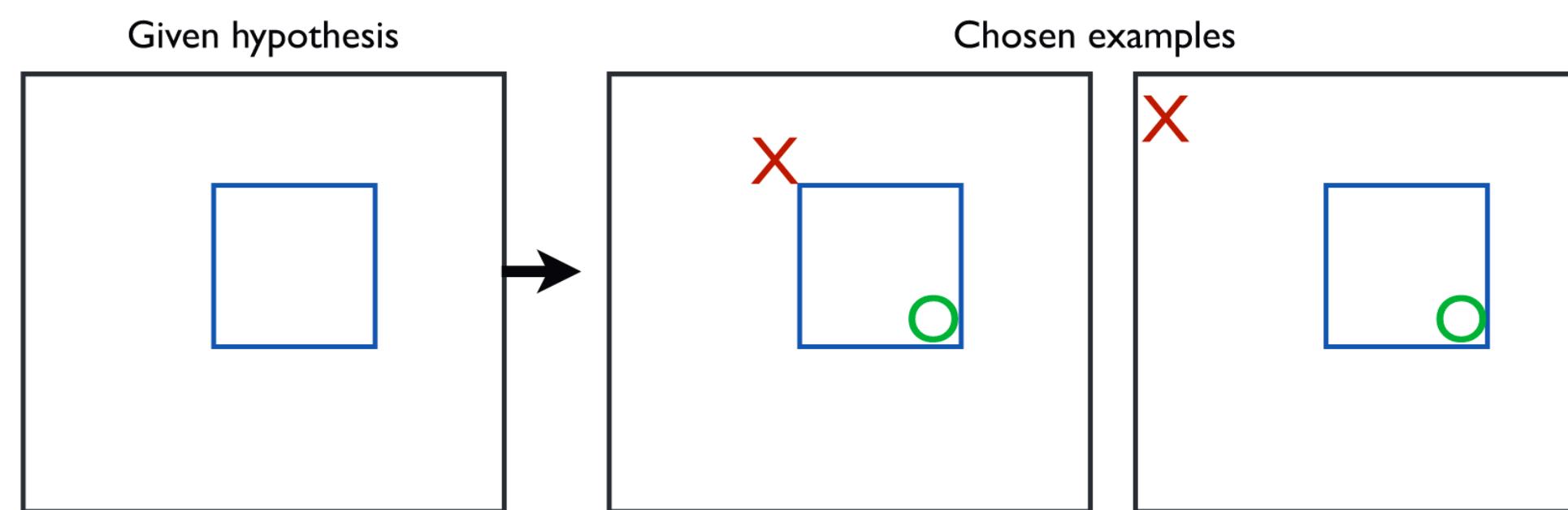


Modeling pedagogical reasoning

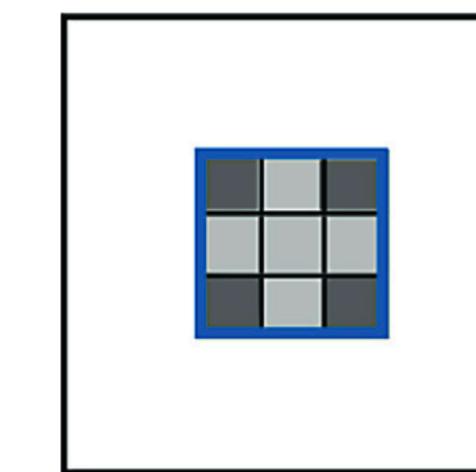
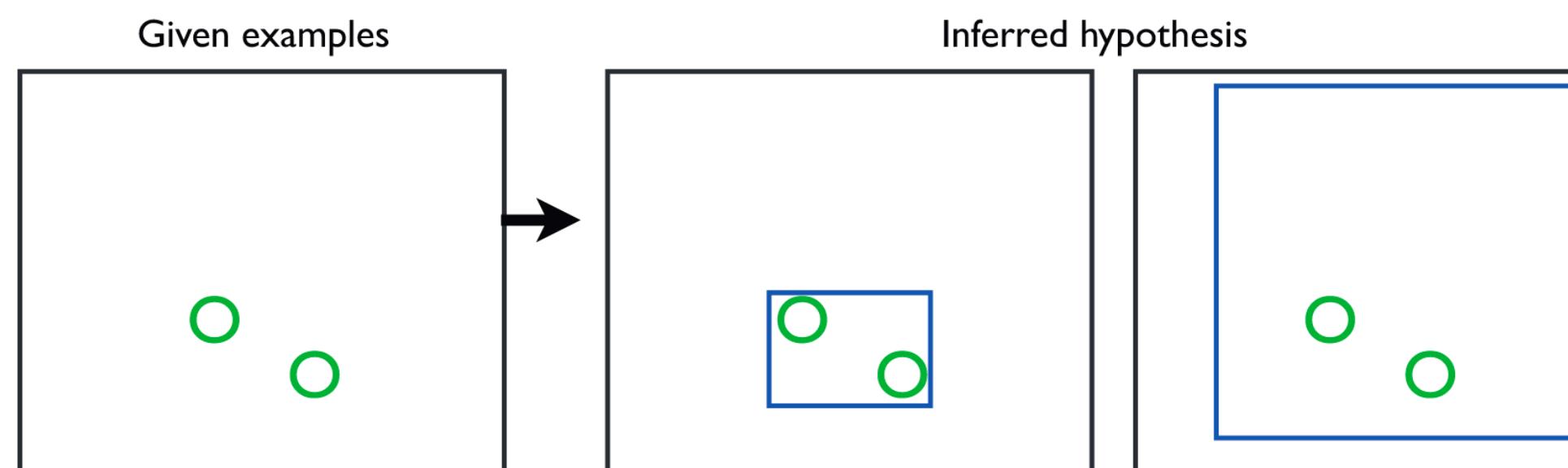
$$P_{\text{teacher}}(d|h) \propto \left(P_{\text{learner}}(h|d)\right)^{\alpha}$$

$$P_{\text{learner}}(h|d) \propto P_{\text{teacher}}(d|h)P(h)$$

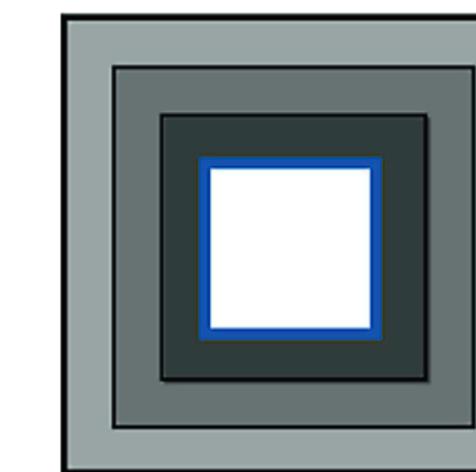
(a) Teacher's perspective



(b) Learner's perspective

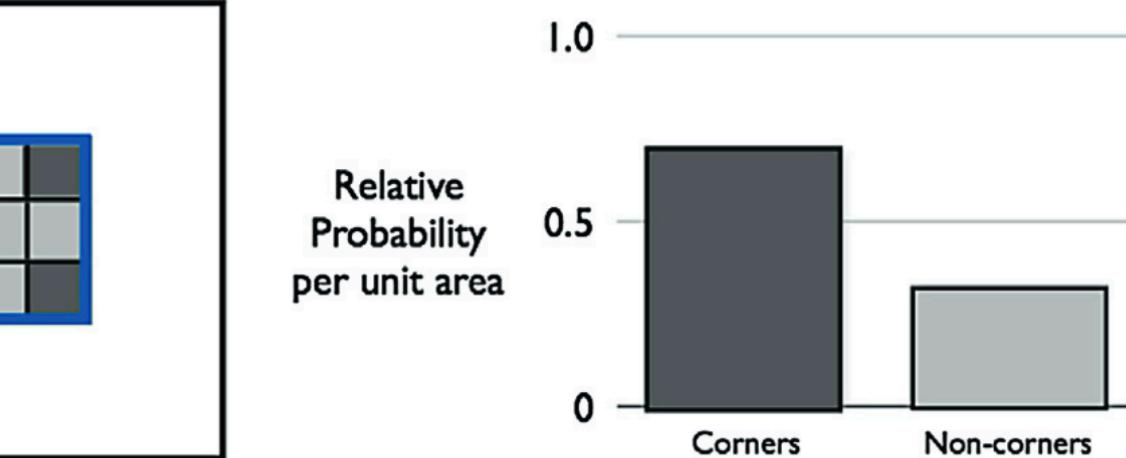


(a) Positive examples.



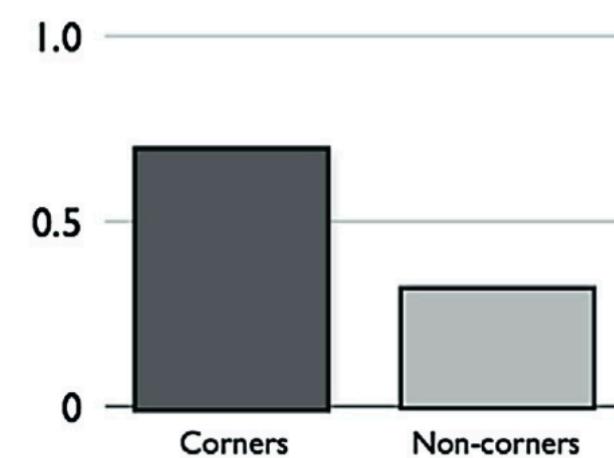
(b) Negative examples.

Relative Proportion per unit area



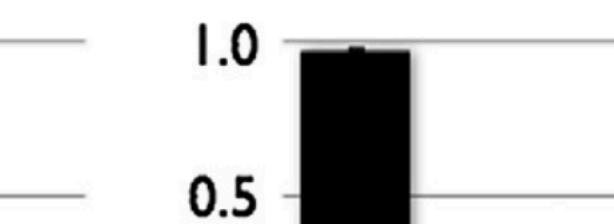
Model

Relative Proportion per unit area

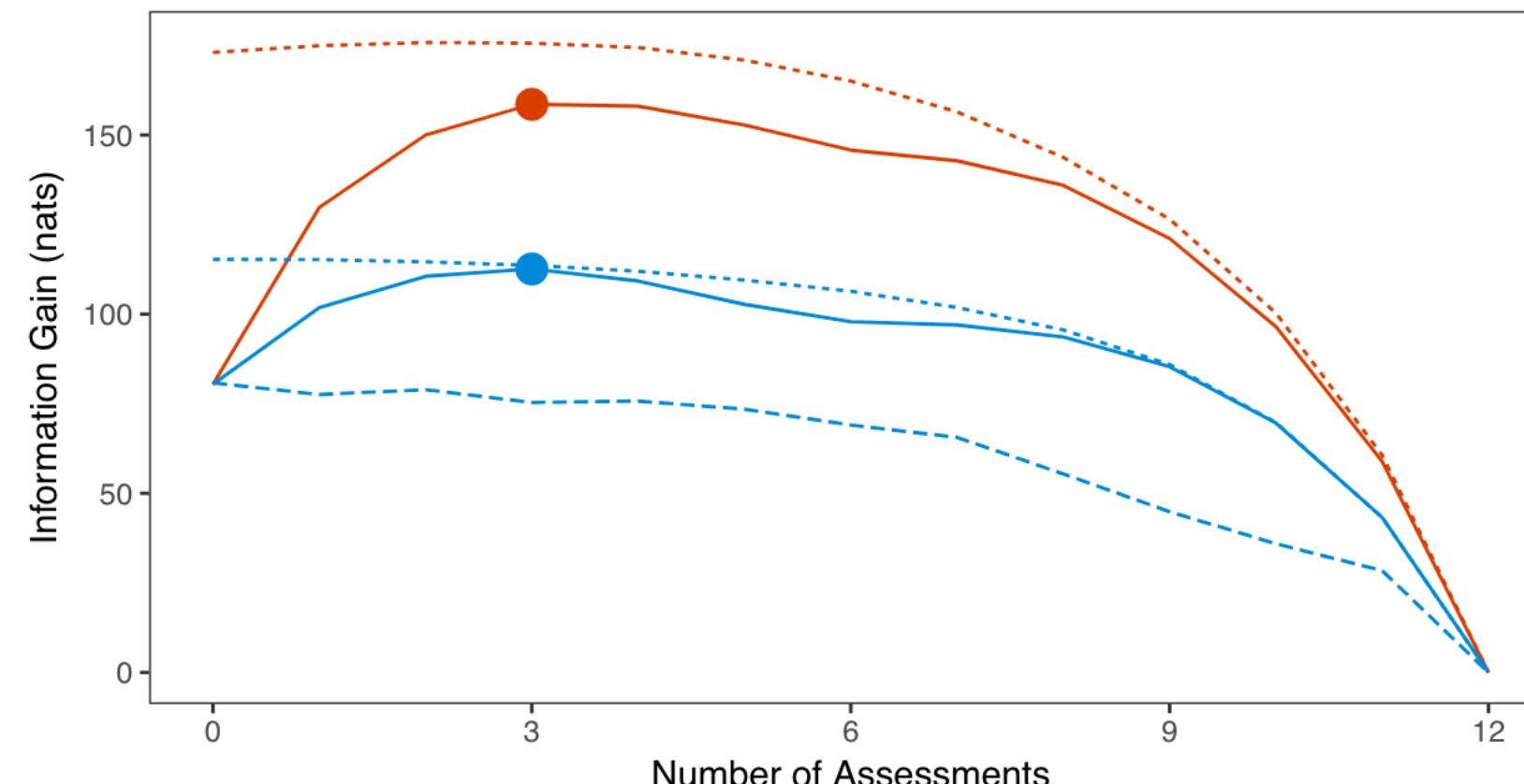


Data

Relative Proportion per unit area

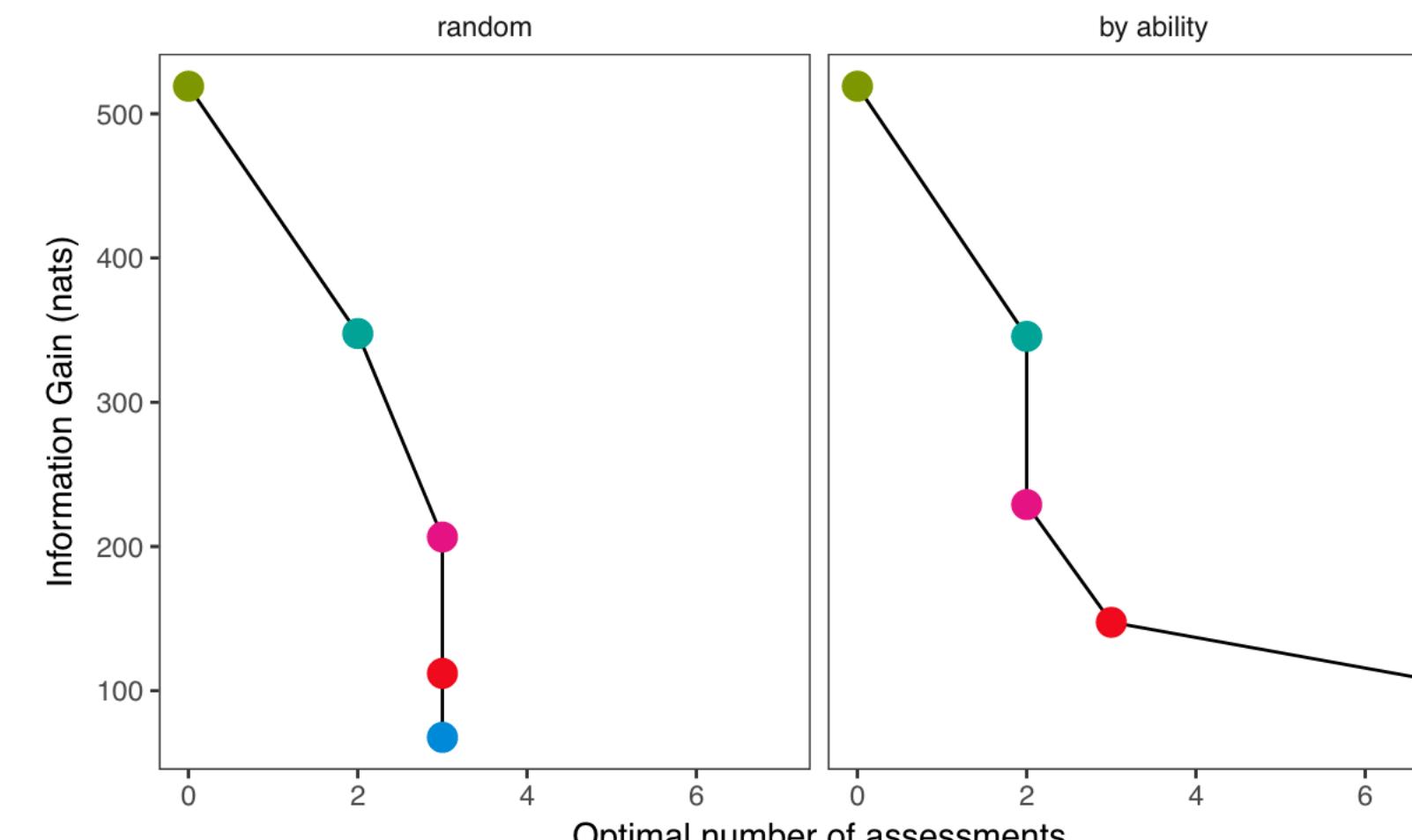
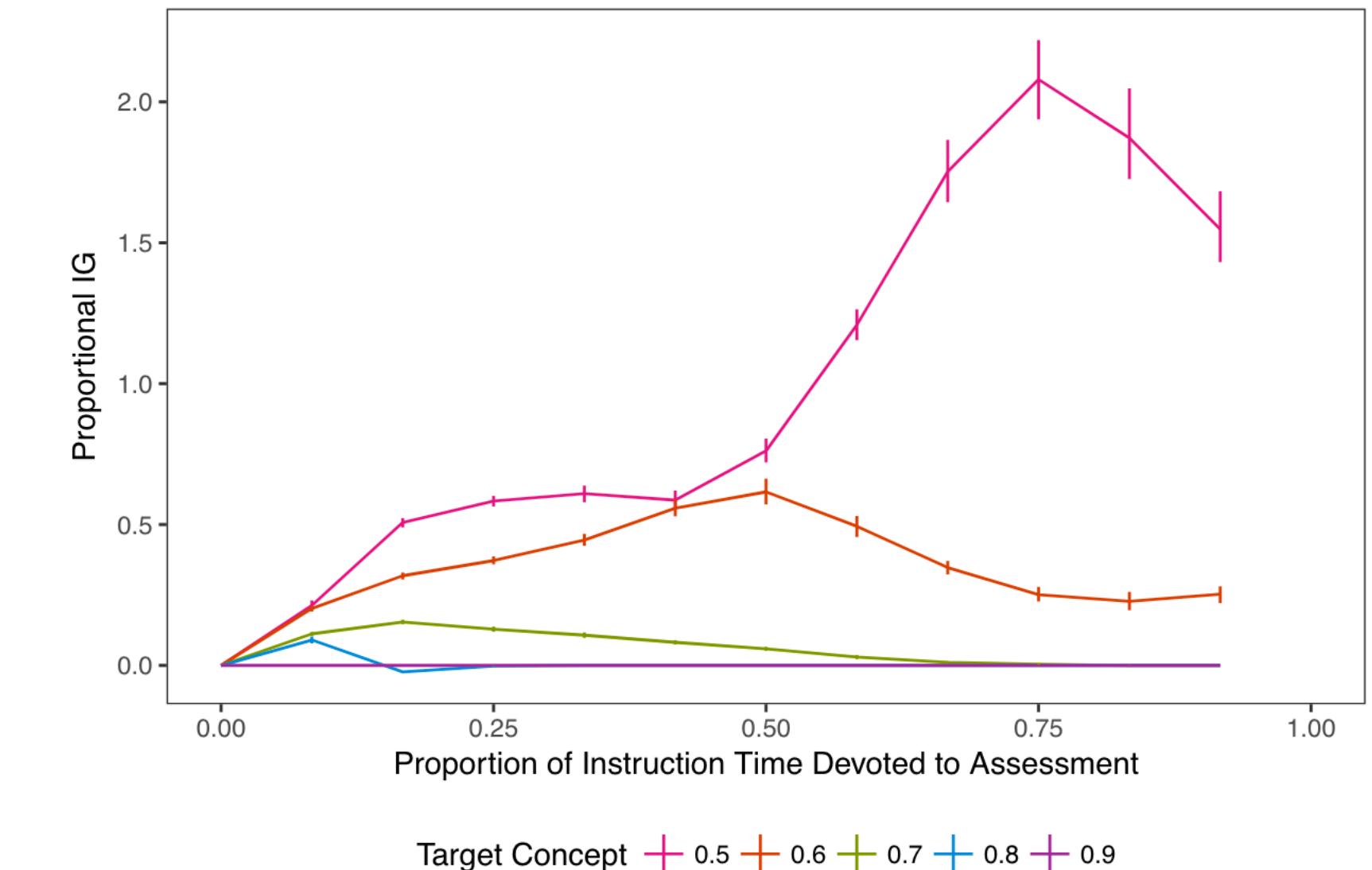


Modeling classroom teaching



Grouping —●— by ability ●—○— random

Teacher Knowledge —— uncertain ---- perfect - - - naive



Target Concept ● 0.5 ● 0.6 ● 0.7 ● 0.8 ● 0.9 ● 0.9

Experiment

Setup

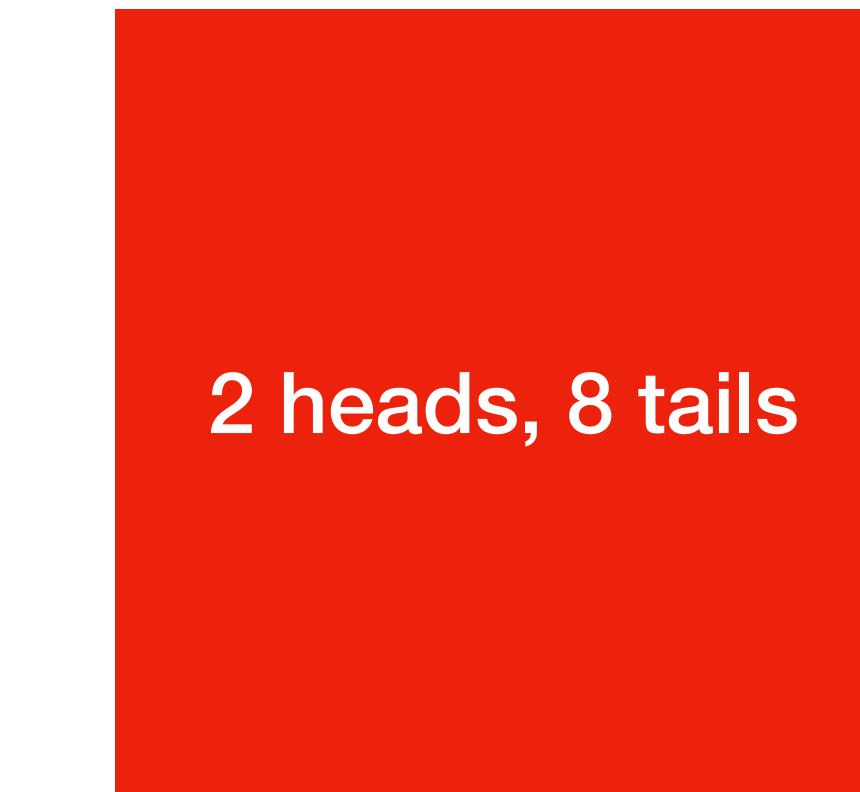
Overview

- Participants are on an alien planet, where coins can be unfair
- They provide **example outcomes** of coin flips to lead students to the correct coin weight
- Flips have a cost
- Participants are rewarded if students' guesses are close to the right coin weight

Setup

Classrooms

- Each student is from **one of four classrooms**, where they have seen a certain # of heads and tails before.



Setup Conditions

- Non-sequential (one set of examples)
 - **full or partial** knowledge of student's classroom
- Sequential (two sets of examples)
 - **feedback or no feedback** from student after first set of examples

Setup

Design

- 120 participants
- 4 conditions
- 3 different coin weights
- 2 possible sets of student classrooms

Sample trial: non sequential + full

Student 6

The weight of your coin is 0.8.

In this block, you will be giving **one** set of examples to Student 6.

You will receive feedback at the end of the block.

[Continue](#)

The weight of your coin is 0.8.

Student info: Student 6 is from Classroom A, where they have seen 1 head and 4 tails.

Select a set of examples to send to your student.

Remember you can only select a max of 30 examples.

1 heads

4 tails

**Sample trial: sequential + partial
+ no feedback**

Student 1

The weight of your coin is 0.2.

In this block, you will be giving **two** sets of examples to Student 1.

You will receive feedback at the end of the block.

Continue

The weight of your coin is 0.2.

Student info: Student 1 is from either classroom A or classroom B.

- The students in Classroom A have already seen **1 head and 4 tails**.
- The students in Classroom B have already seen **4 heads and 1 tail**.

Select a set of examples to send to your student.

Remember you can only select a max of 30 examples.

Select another set of examples to send to the same student.

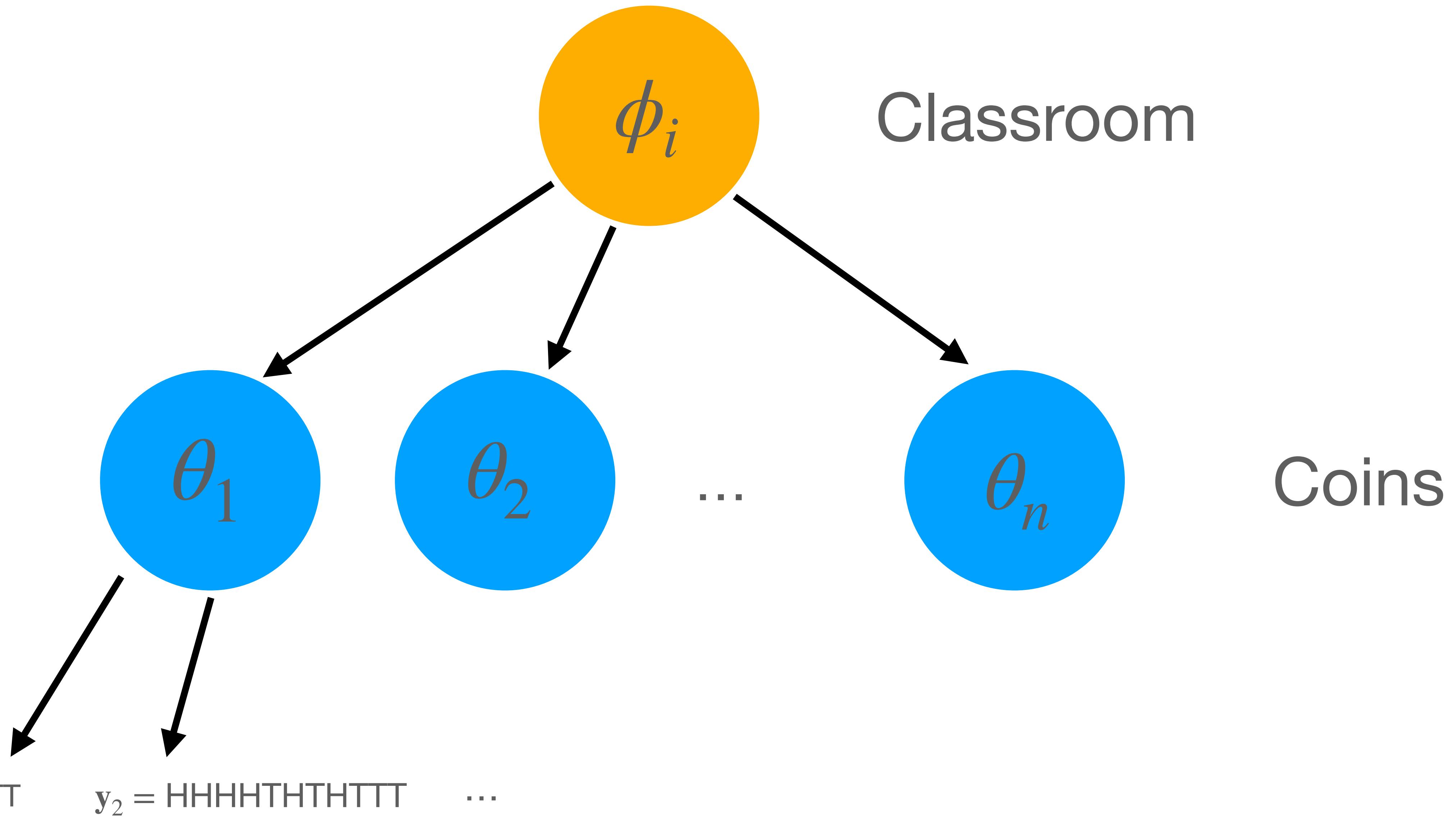
Reminders:

- The weight of your coin is 0.2.
- Student 1 is from either Classroom A or Classroom B.
- The students in Classroom A have already seen **1 head and 4 tails**.
- The students in Classroom B have already seen **4 heads and 1 tail**.

Remember you can only select a max of 30 examples.

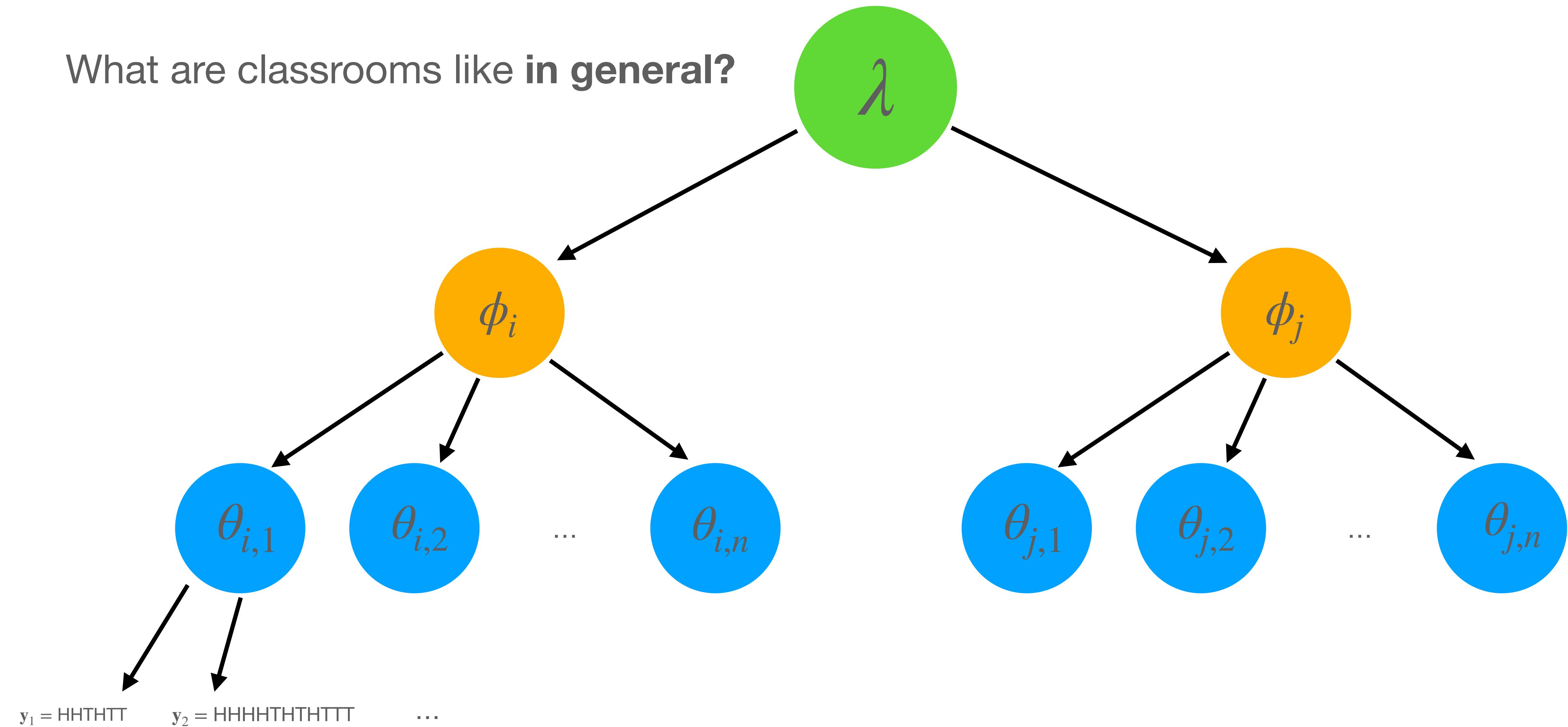
  heads  tails

Model

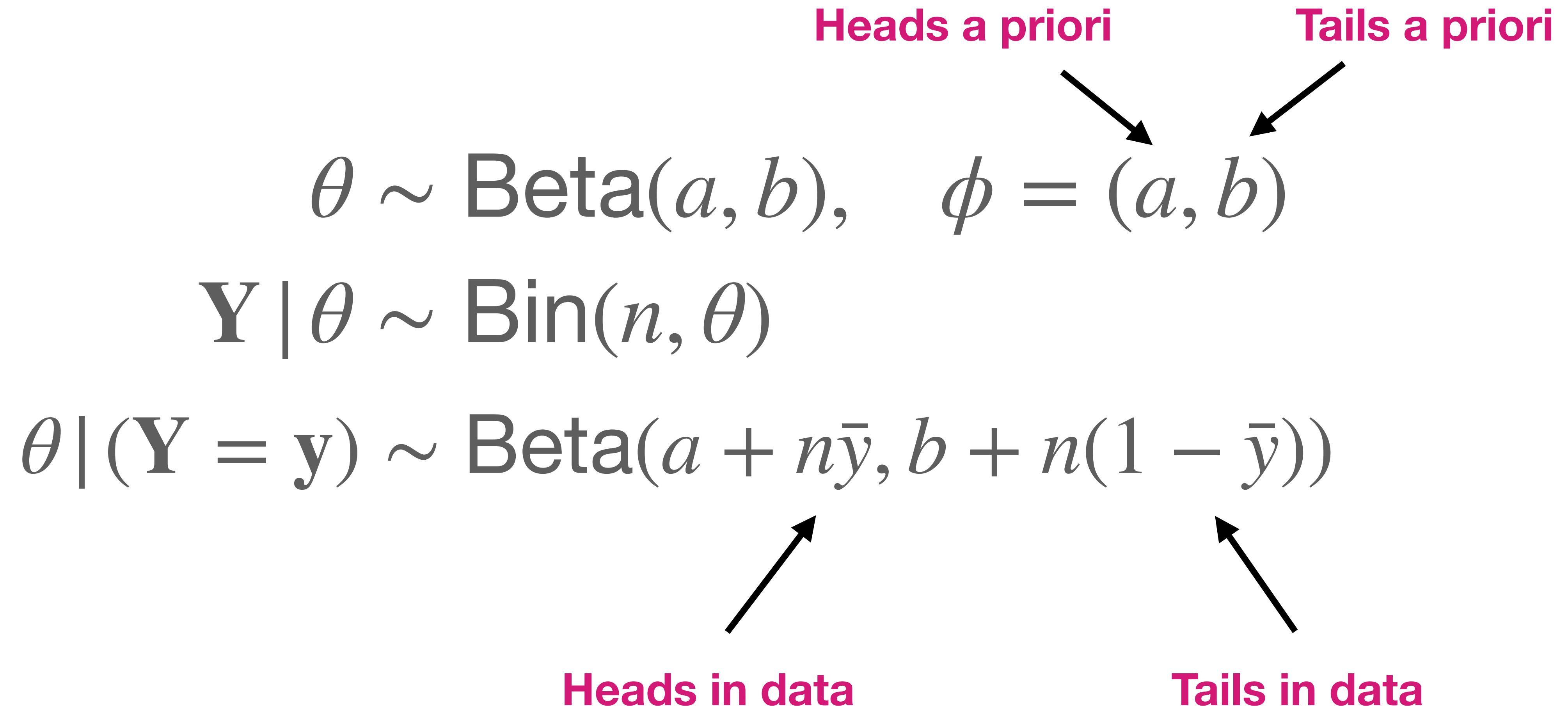


Examples

What are classrooms like in general?



Full knowledge of student's classroom



Full knowledge of student's classroom

$$L_0(\theta | \mathbf{Y}) \propto P(\mathbf{Y} | \theta)P(\theta)$$

Literal
learner

$$U(\mathbf{Y}; \theta) = \log L_0(\theta, \mathbf{Y}) - w_c \cdot C(\mathbf{Y})$$

$$S_1(\mathbf{Y} | \theta) \propto \exp(\alpha \cdot U(\mathbf{Y}; \theta))$$

Pragmatic
speaker

Rational Speech Act models of
pragmatic reasoning; Frank &
Goodman (2016)

Partial knowledge of student's classroom

$$\phi_i = (a_i, b_i)$$

$$\phi \sim \text{Categorical}([\phi_1, \dots, \phi_n])$$

$$\theta | \phi \sim \text{Beta}(a_i, b_i)$$

$$\mathbf{Y} | \theta, \phi \sim \text{Bin}(n, \theta)$$

$$L_0(\theta | \mathbf{Y}, \phi) \propto P(\mathbf{Y} | \theta, \phi)P(\theta | \phi)$$

$$U(\mathbf{Y}; \theta, \phi) = \log L_0(\theta | \mathbf{Y}, \phi) - w_c \cdot C(\mathbf{Y})$$

$$S_1(\mathbf{Y} | \theta, \phi) \propto \exp(\alpha \cdot U(\mathbf{Y}; \theta, \phi))$$

Integrate over
possible
classrooms

$$S(\mathbf{Y} | \theta) = \sum_{\phi} S_1(\mathbf{Y} | \theta, \phi)P(\phi)$$

Sequential example selection

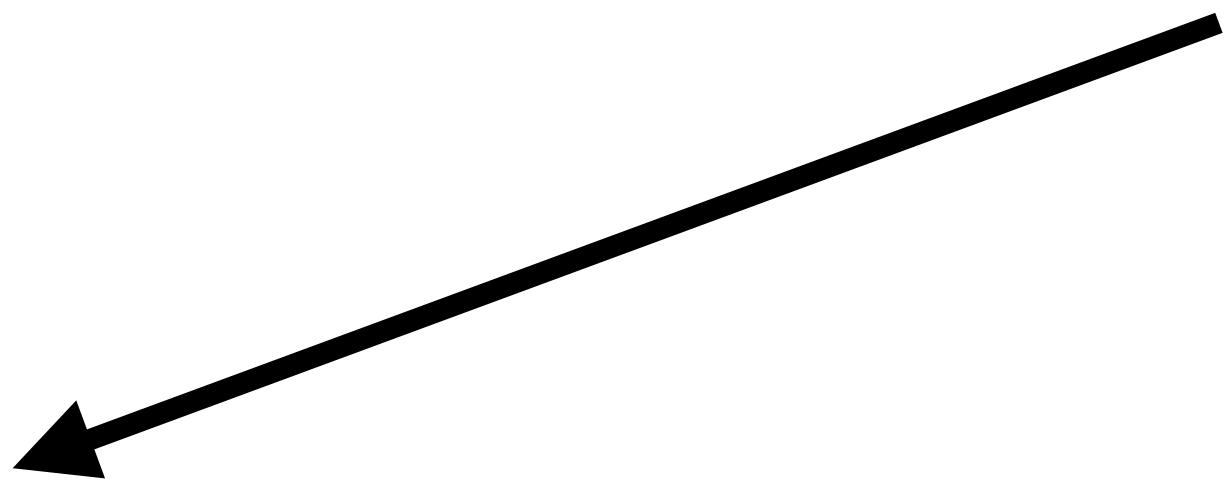
No feedback

$$\begin{aligned} S(\mathbf{Y}_2 \mid \mathbf{y}_1, \theta) &= \sum_{\phi} S_1(\mathbf{Y}_2 \mid \theta, \phi) P(\phi \mid \mathbf{y}_1, \theta) \\ &= S(\mathbf{Y} \mid \theta) \end{aligned}$$

Sequential example selection

With feedback

$$S(\mathbf{Y}_2, \mathbf{Y}_1 | \theta) = S(\mathbf{Y}_2 | \mathbf{y}_1, \theta)S(\mathbf{Y}_1 | \theta)$$



$$S(\mathbf{Y}_2 | \mathbf{y}_1, \theta) = \int_{\hat{\theta}_L} S(\mathbf{Y}_2 | \hat{\theta}_L, \theta)P(\hat{\theta}_L | \mathbf{y}_1, \theta)$$

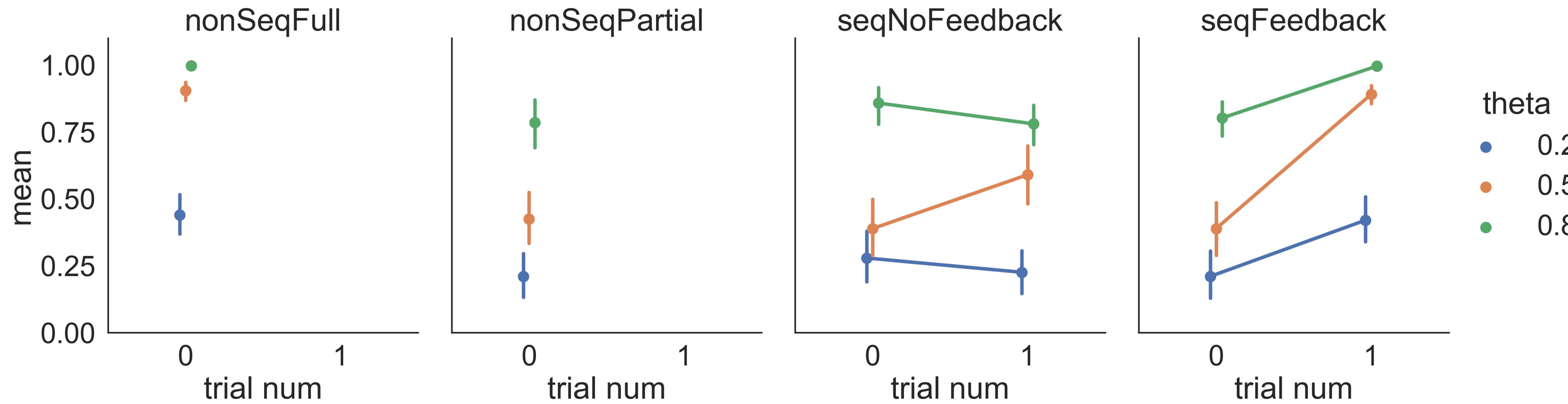
$$S(\mathbf{Y}_2 | \hat{\theta}_L, \theta) \propto \sum_{\phi} S_1(\mathbf{Y}_2 | \phi, \hat{\theta}_L, \theta)P(\phi | \hat{\theta}_L, \theta)$$

$$P(\phi | \hat{\theta}_L, \theta) \propto P(\hat{\theta}_L | \phi, \theta)P(\phi | \theta)$$

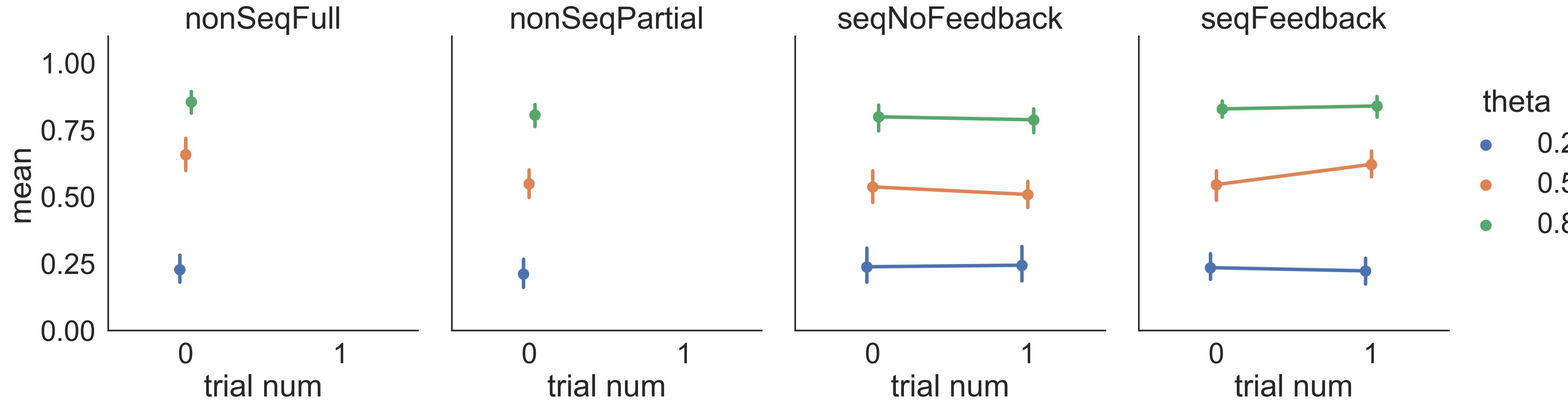
$$\hat{\theta}_L = \mathbb{E}[\theta | \mathbf{y}_1, \phi^*]$$

Results

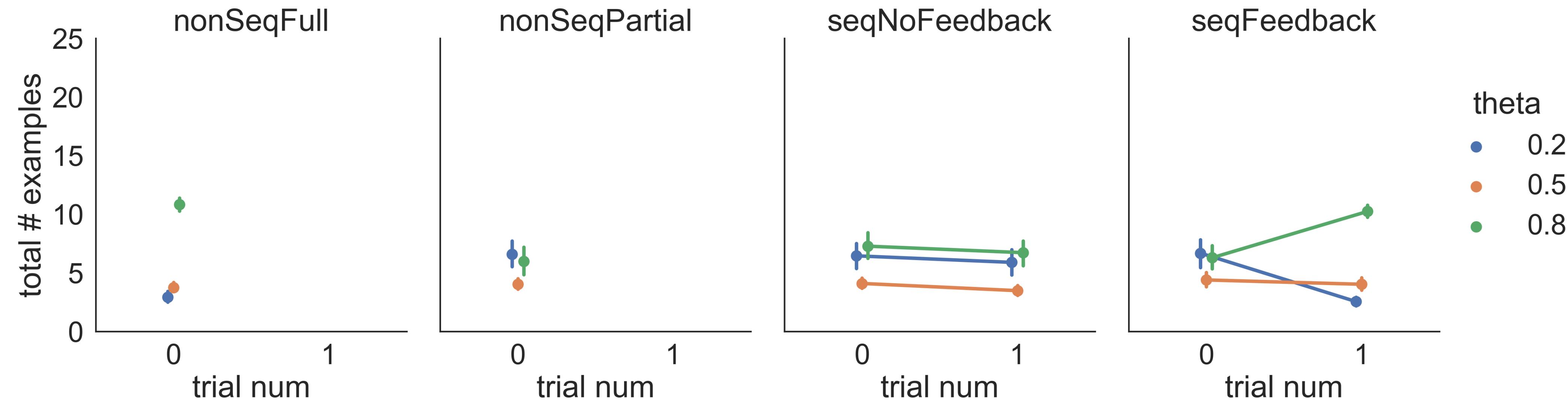
model ($a = 1, b = 4$) proportion heads



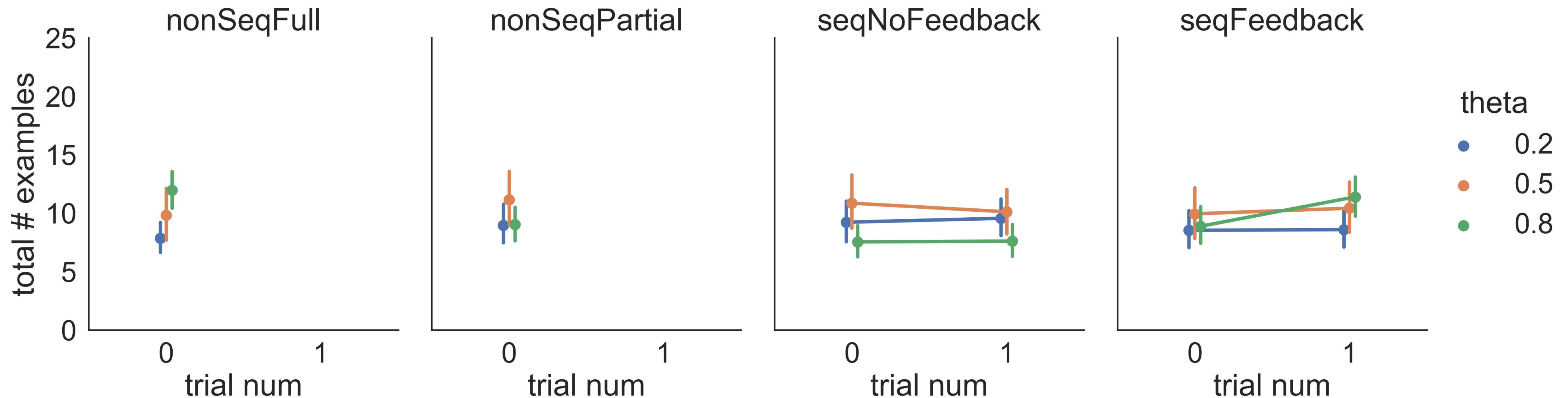
data ($a = 1, b = 4$) proportion heads



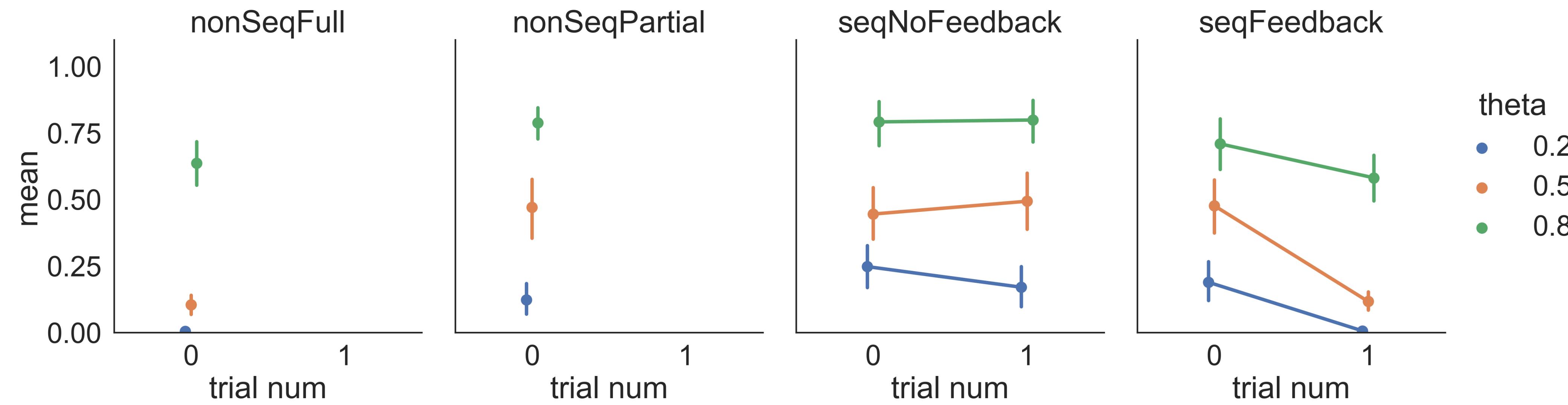
model ($a = 1, b = 4$) total # examples



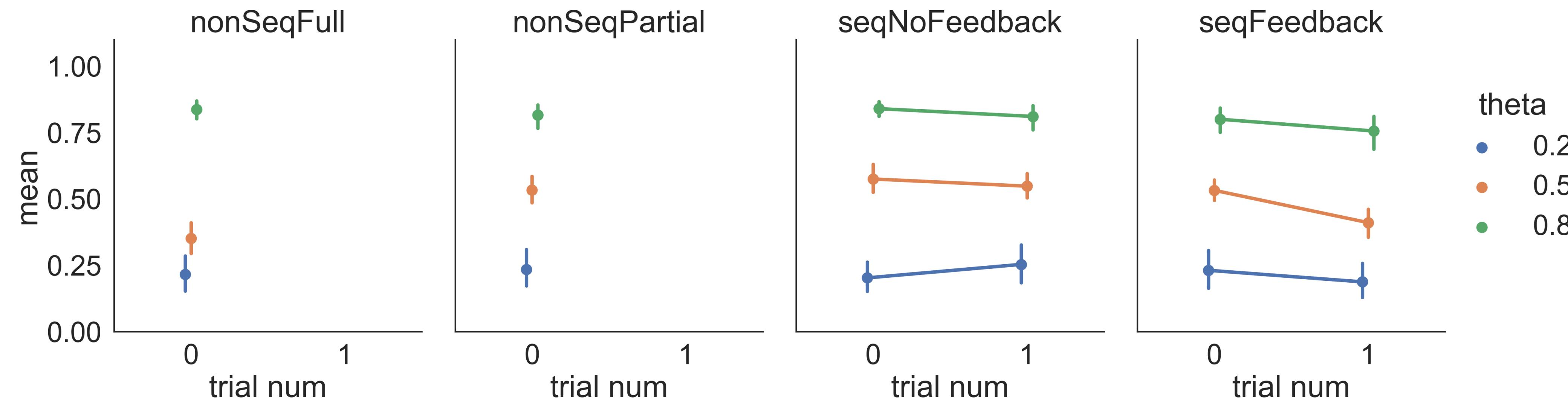
data ($a = 1, b = 4$) total # examples



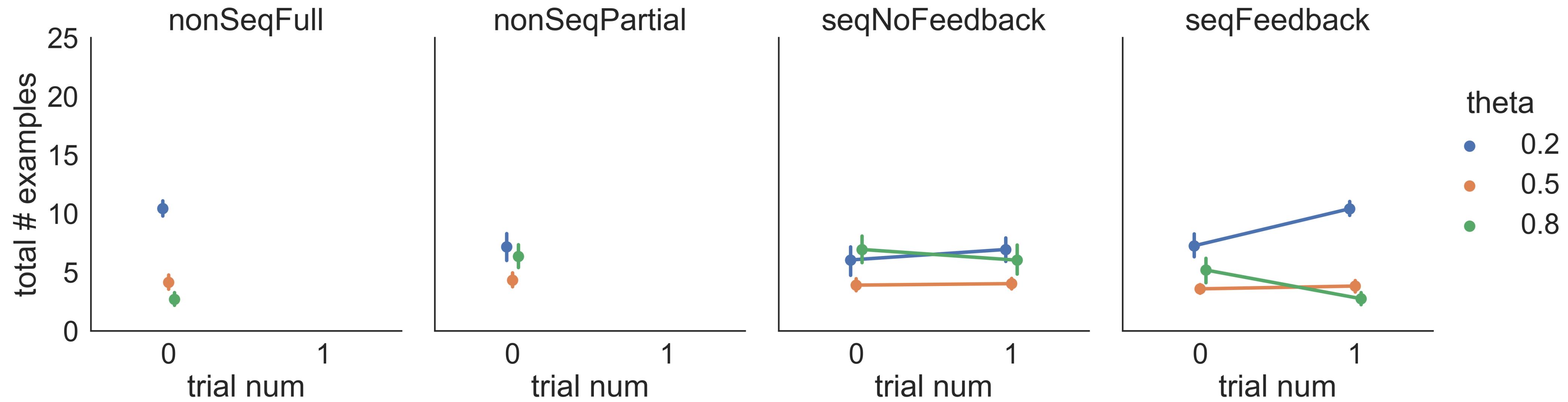
model ($a = 4, b = 1$) proportion heads



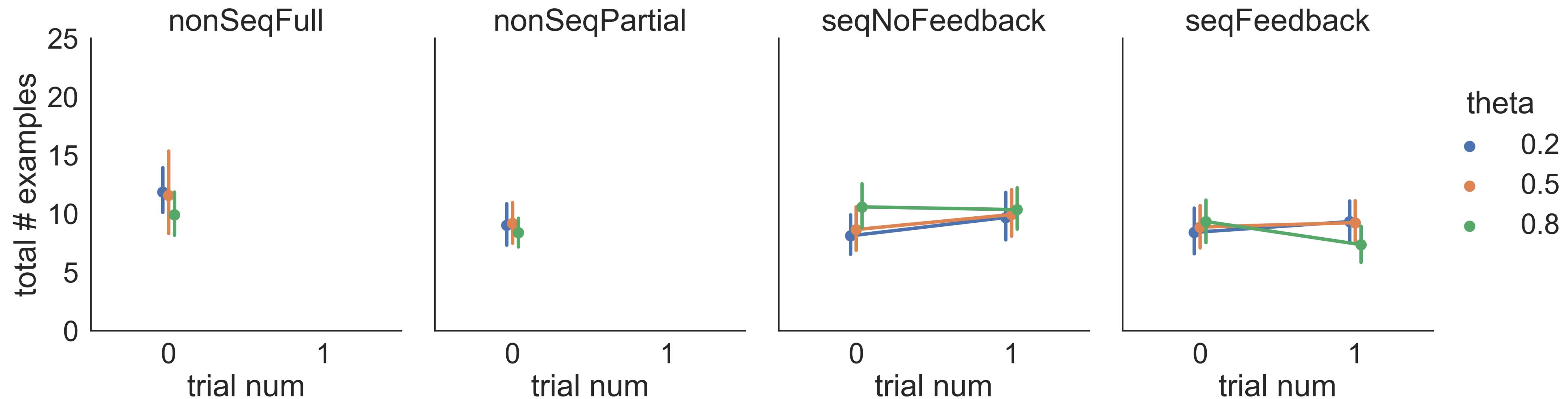
data ($a = 4, b = 1$) proportion heads



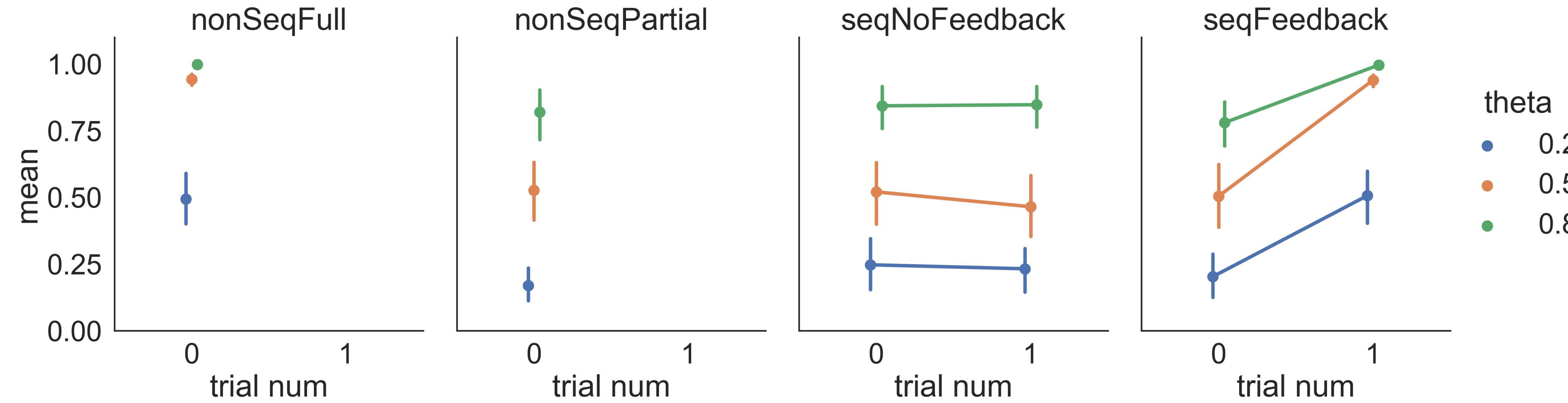
model ($a = 4, b = 1$) total # examples



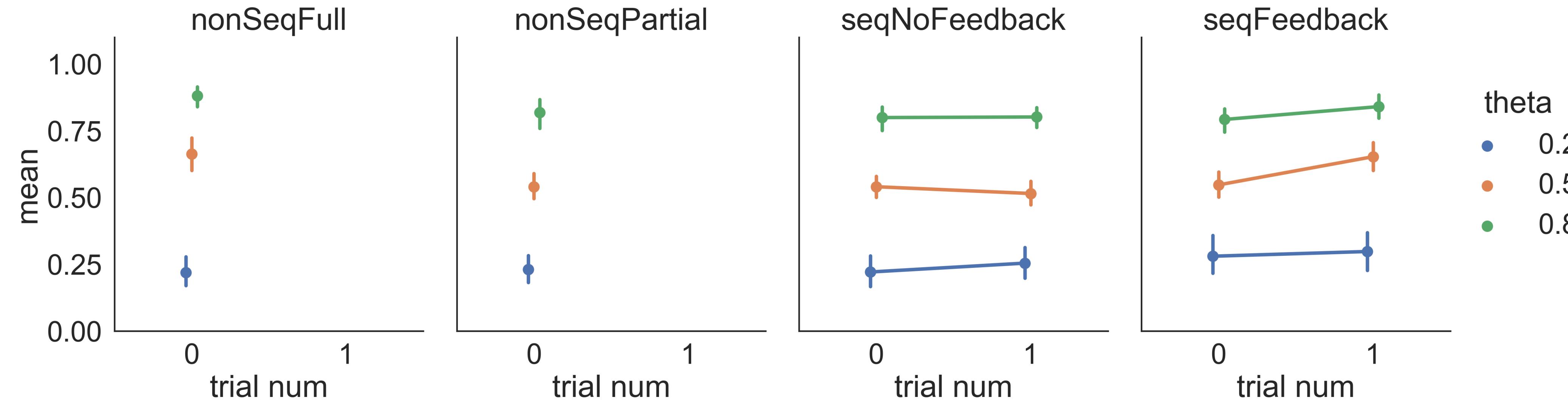
data ($a = 4, b = 1$) total # examples



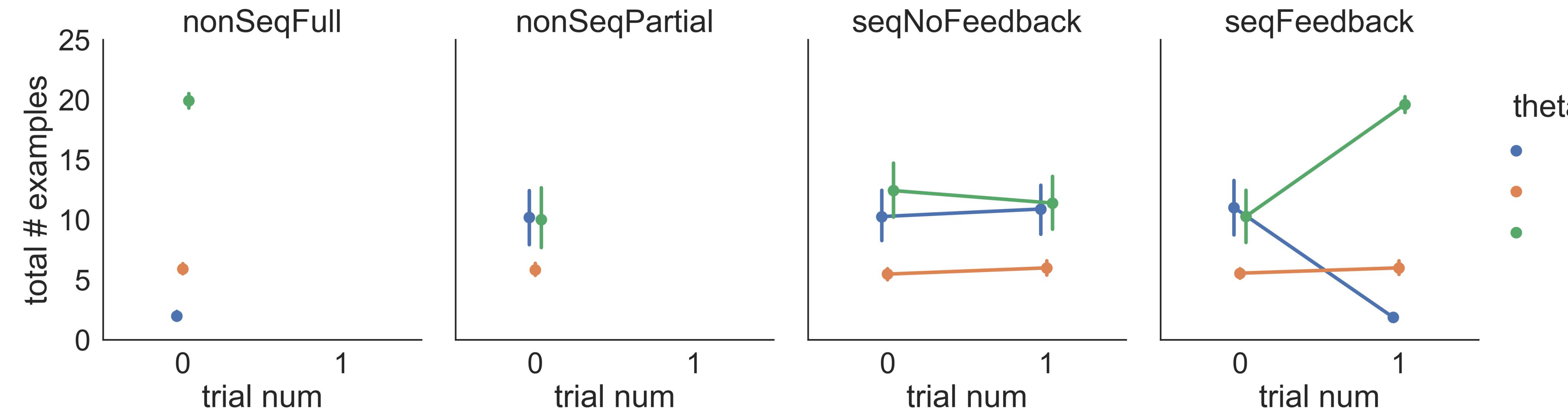
model ($a = 2, b = 8$) proportion heads



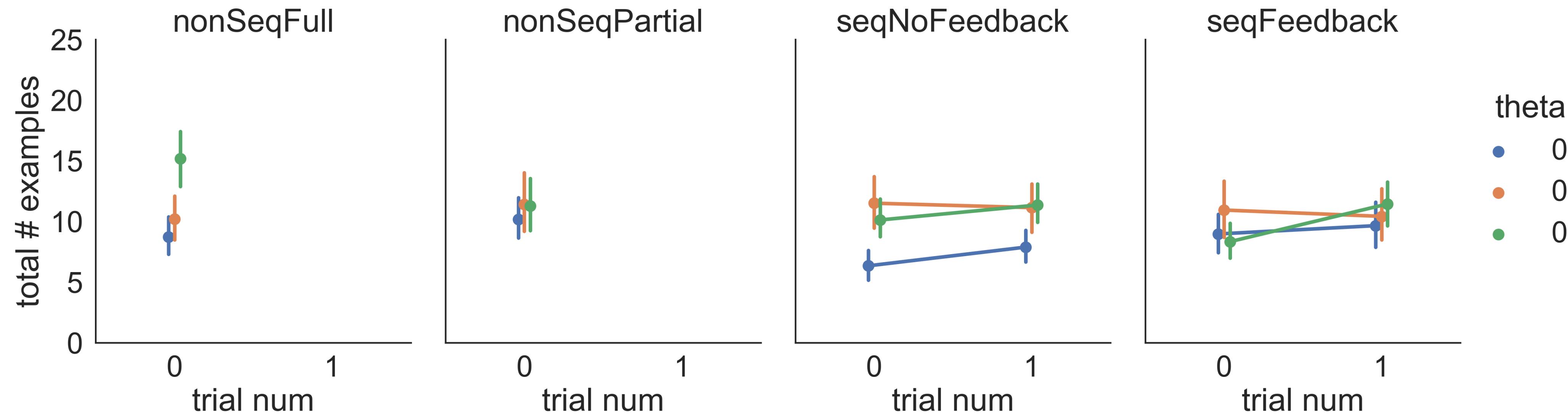
data ($a = 2, b = 8$) proportion heads



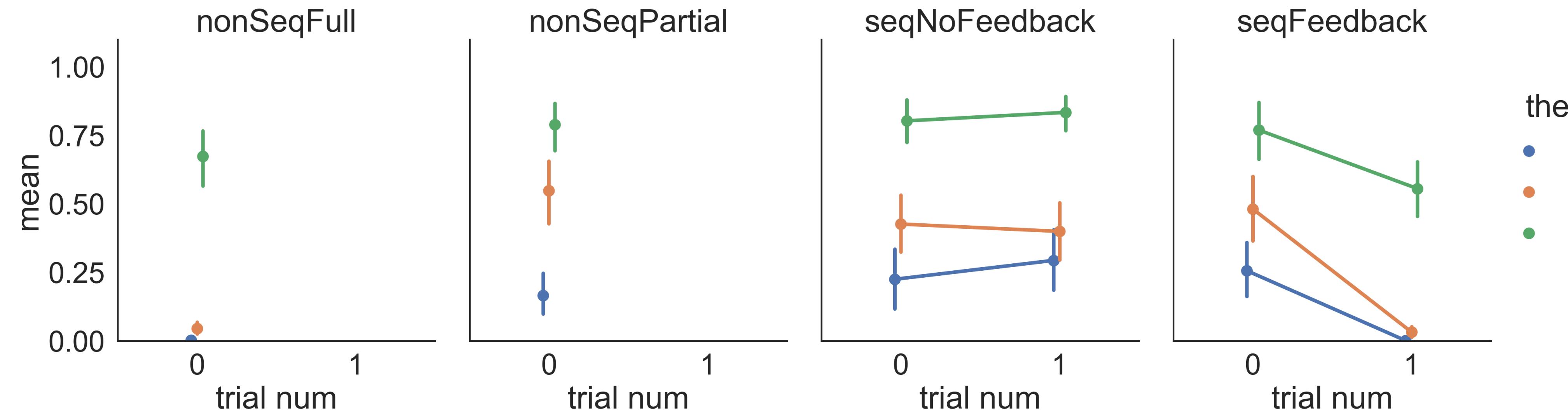
model ($a = 2, b = 8$) total # examples



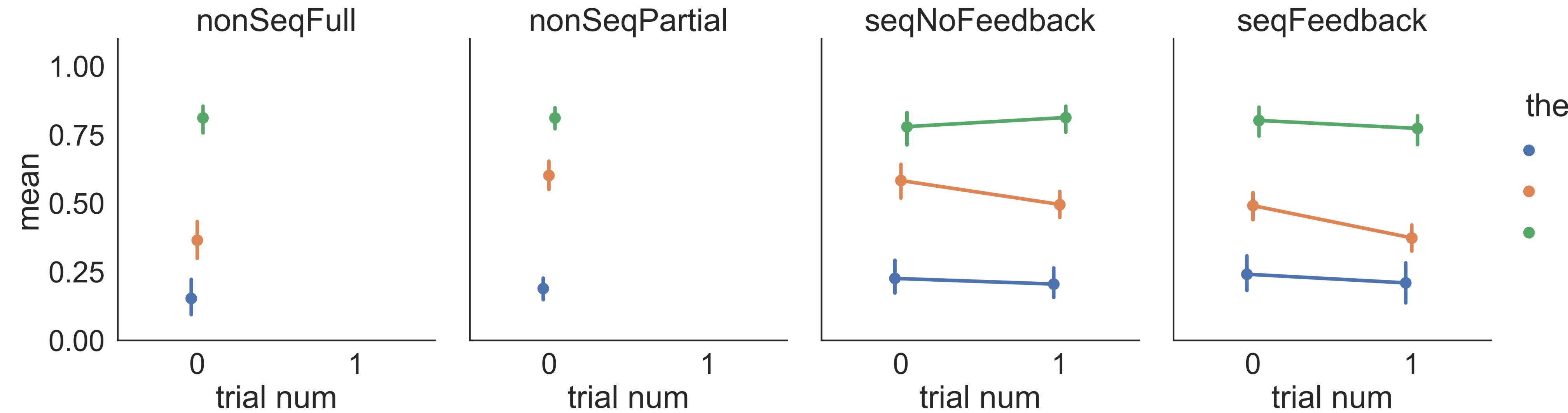
data ($a = 2, b = 8$) total # examples



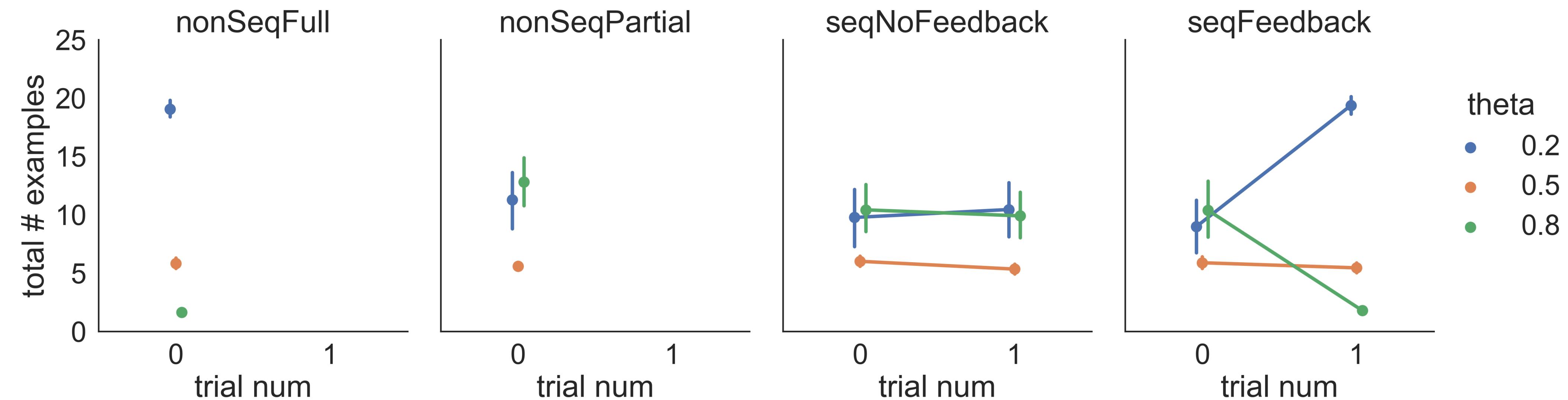
model ($a = 8, b = 2$) proportion heads



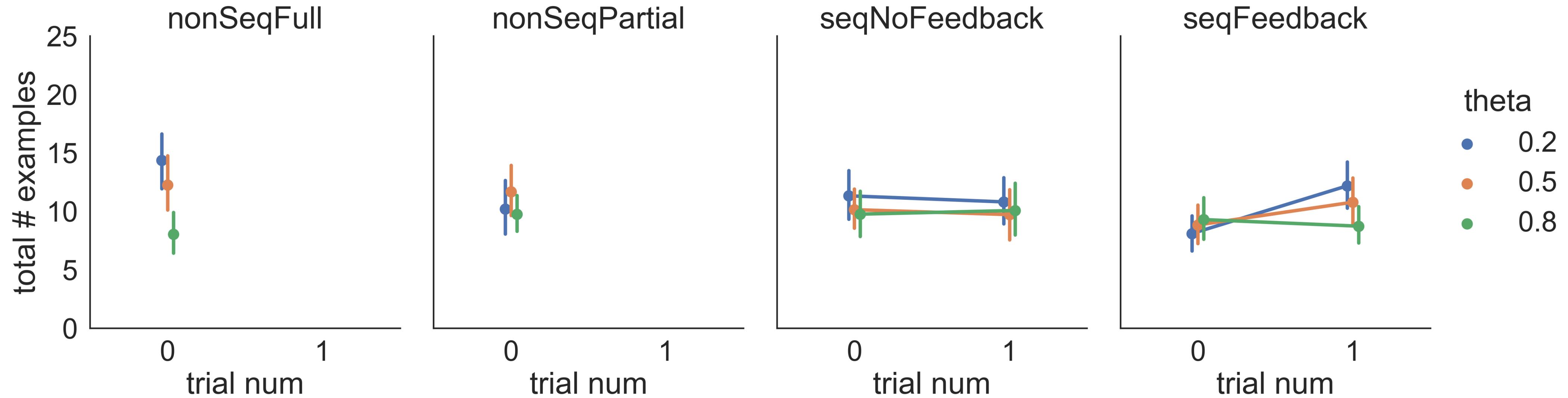
data ($a = 8, b = 2$) proportion heads



model ($a = 8, b = 2$) total # examples



data ($a = 8, b = 2$) total # examples



Issues

- Coin flip setup is confusing to participants
 - Alternative backstories
- Form of feedback
- More layers of recursion? Teacher teaching to a pragmatic listener