Item Ordering Biases in Educational Data

Jaroslav Čechák Radek Pelánek

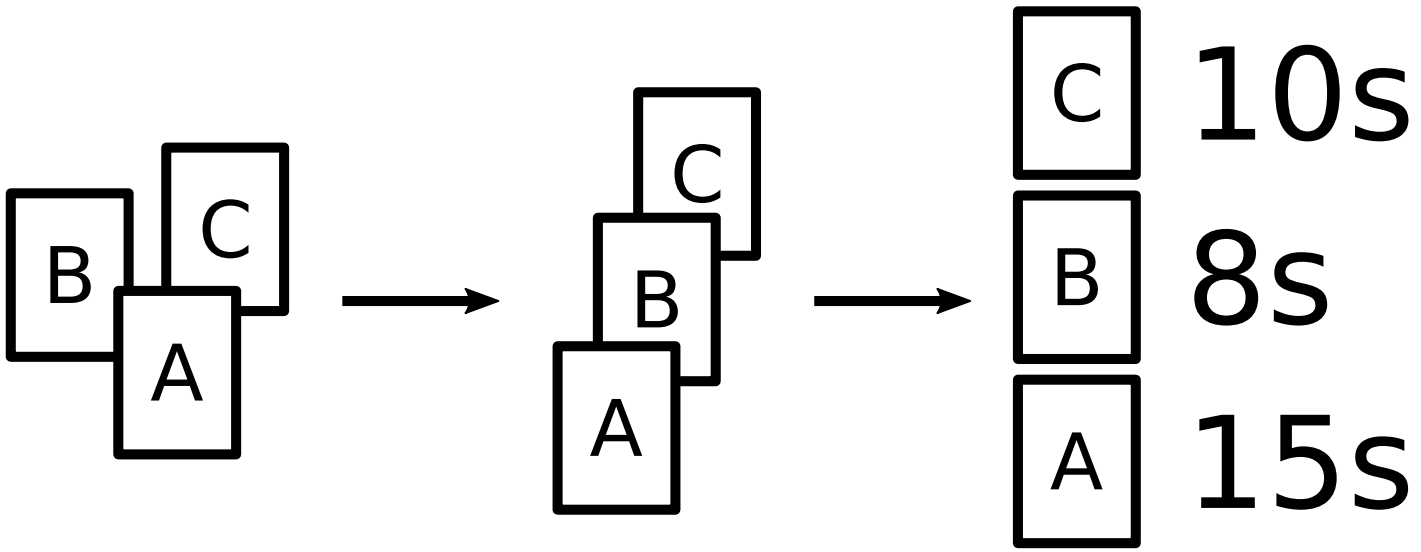
Adaptive Learning
Research group, Masaryk university Brno

AIED 2019

Biases in Educational Data

- self-selection
- attrition, mastery
- **item ordering**

Item Ordering Bias Example



Why Does It Matter?

- collected data are used to modify the system
- bias in data $\rightarrow$ misleading conclusion $\rightarrow$ wrong action

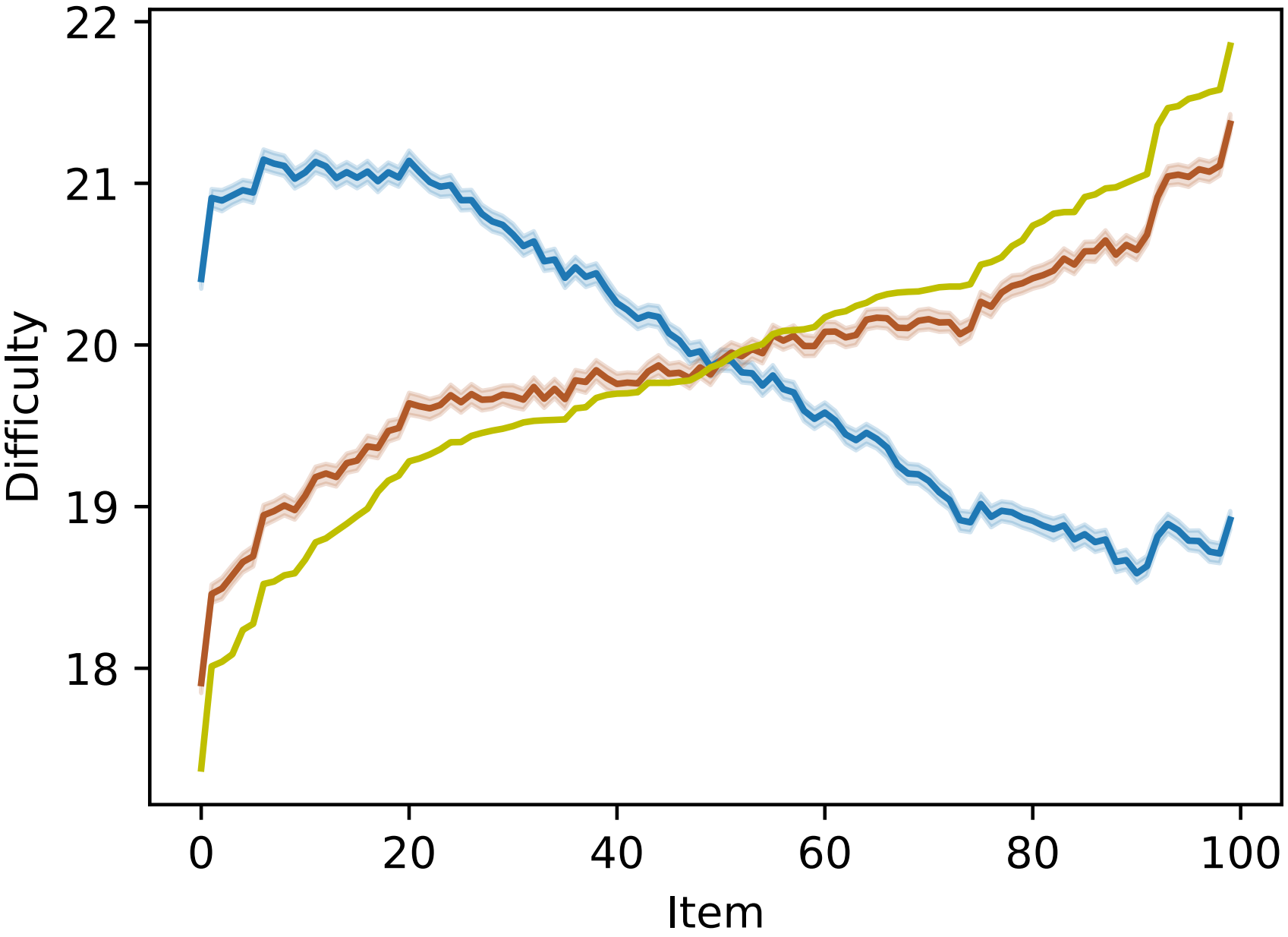
How to Explore It?

- difficult from already collected data
- easy through simulations

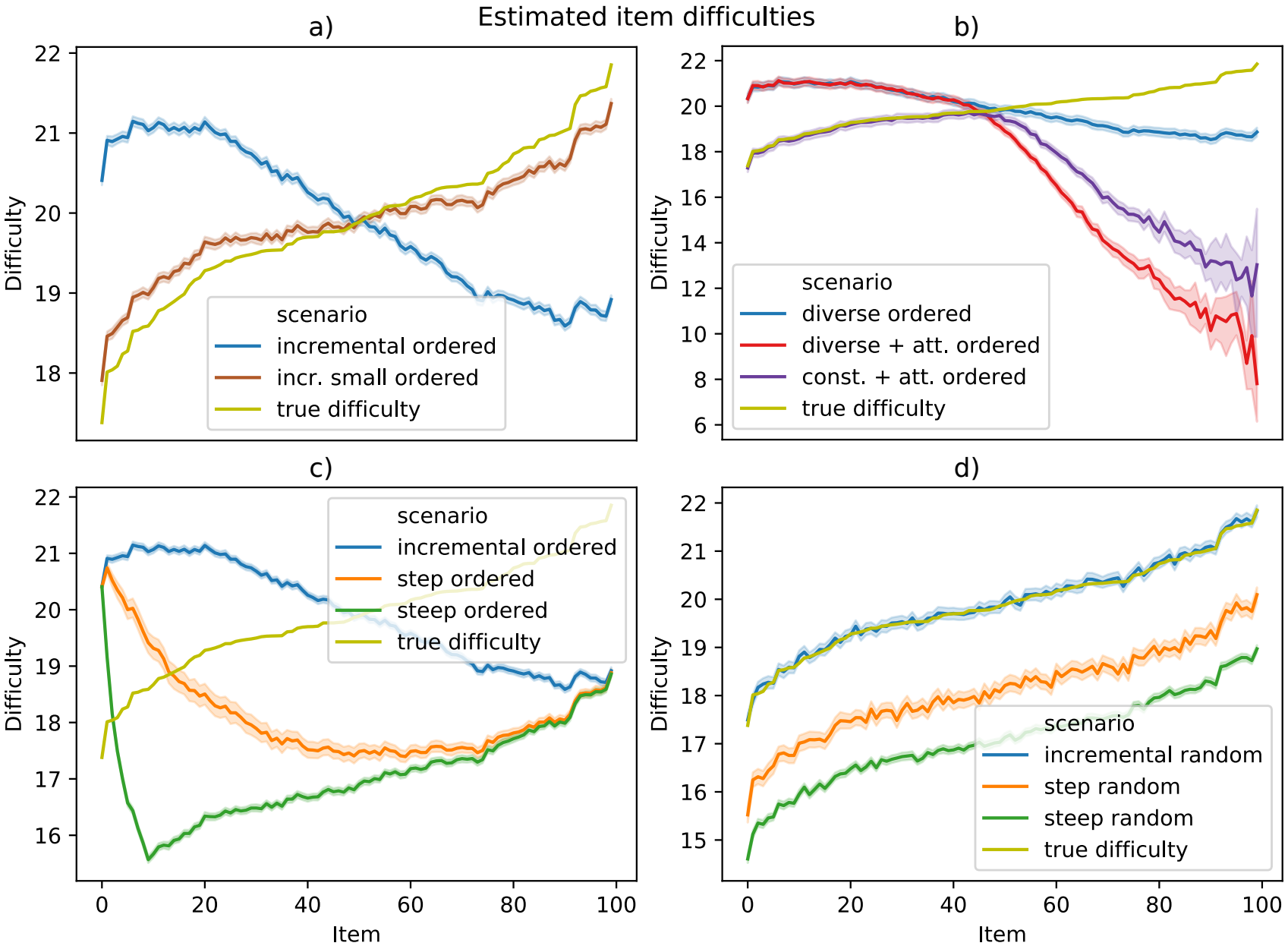
Simulation Setup

1. take an available student
2. select an item
3. generate a solving time using a model
4. repeat

Simulation Results



Simulation Results (cont.)



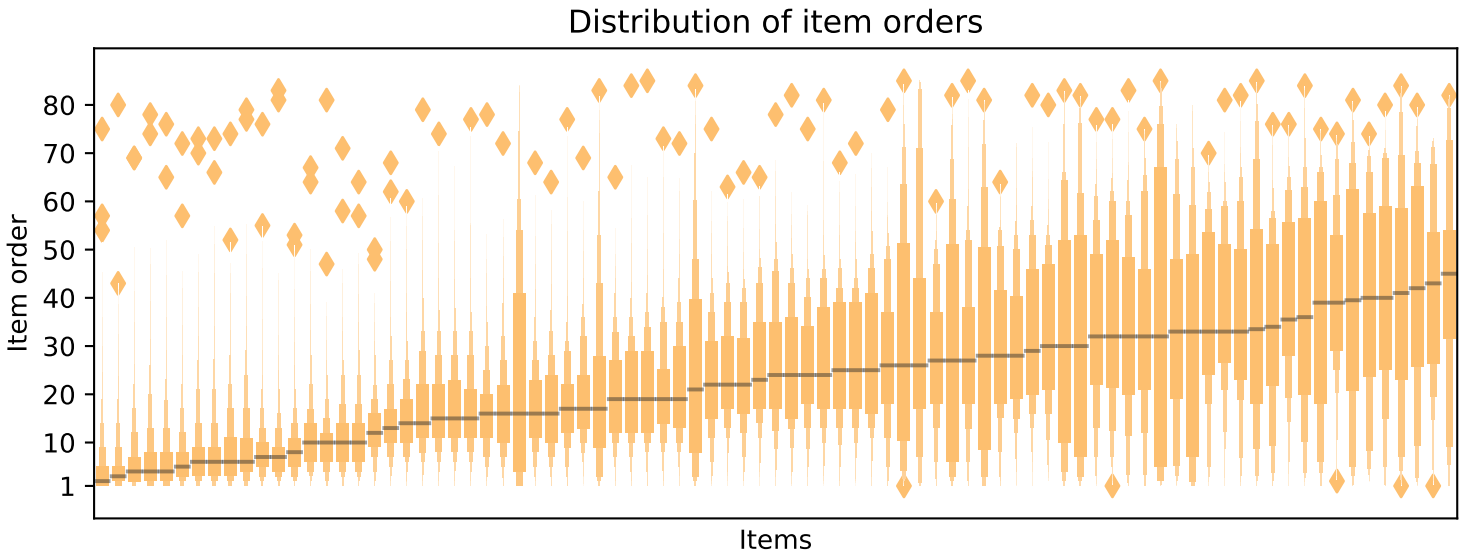
Simulation Results – Takeaway

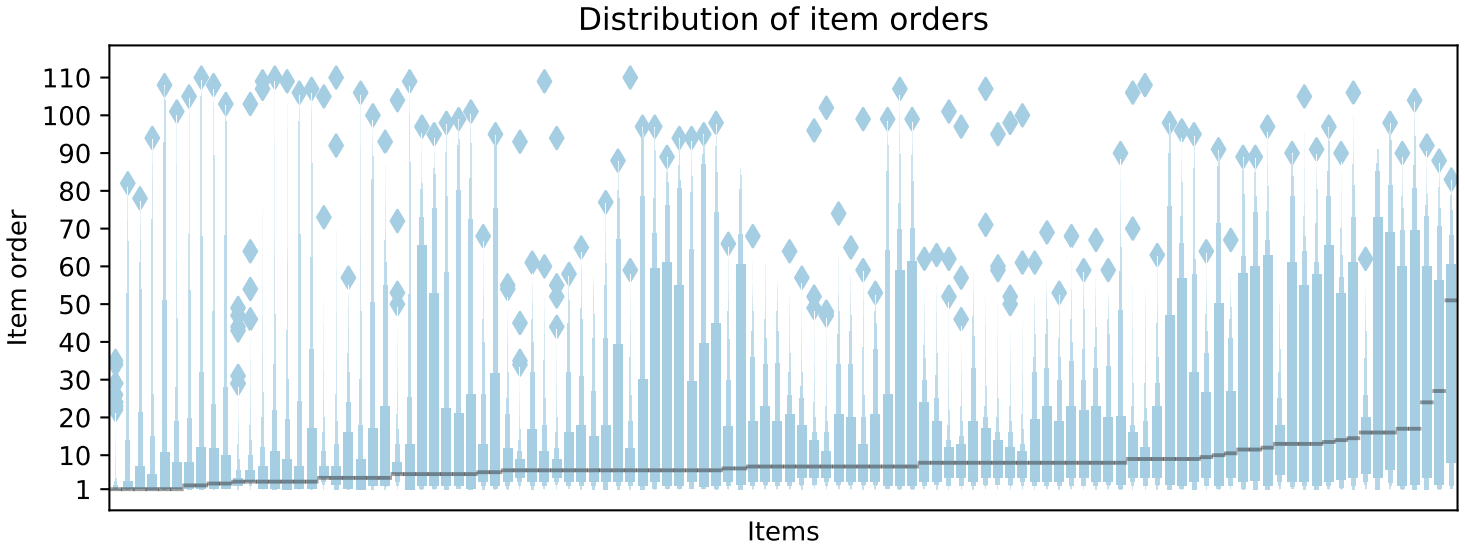
- relation between skill and difficulty increases is important
- randomization helps
- attrition can amplify the problem

Is It Relevant to Me?

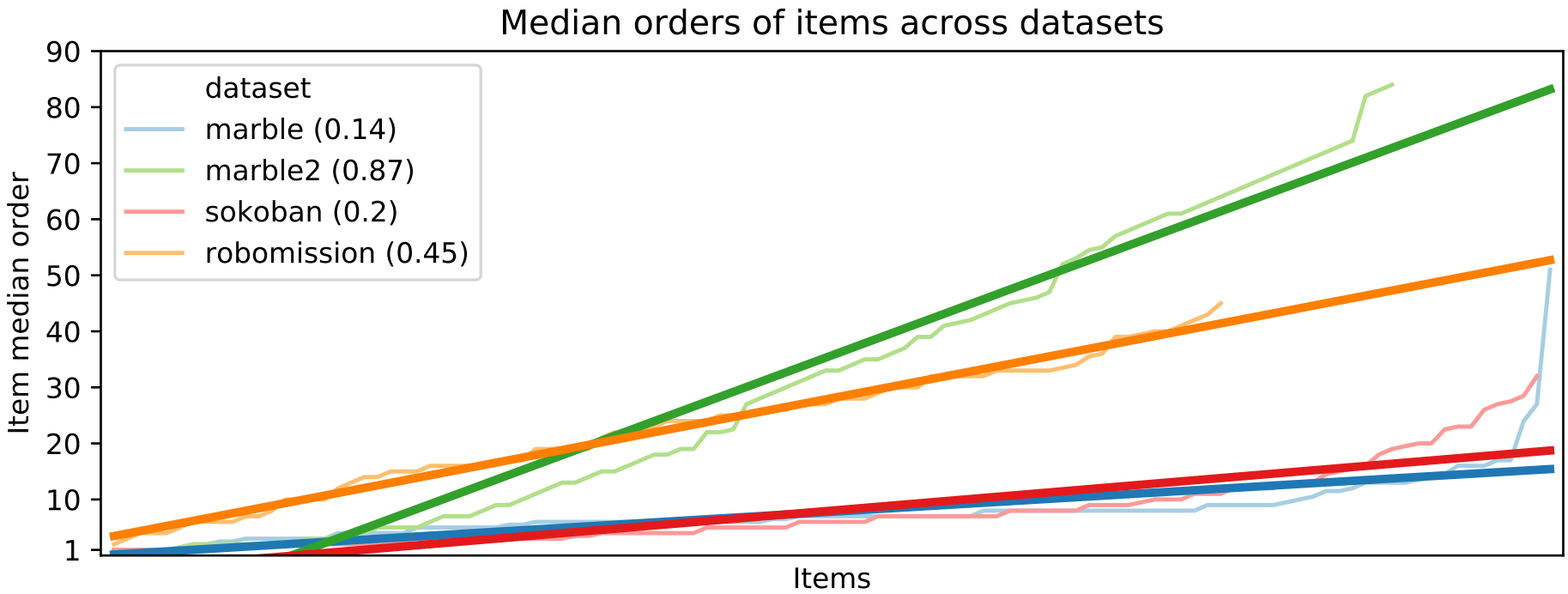
- it depends
- are student solving sequences random

Is It Relevant to Me? (cont.)

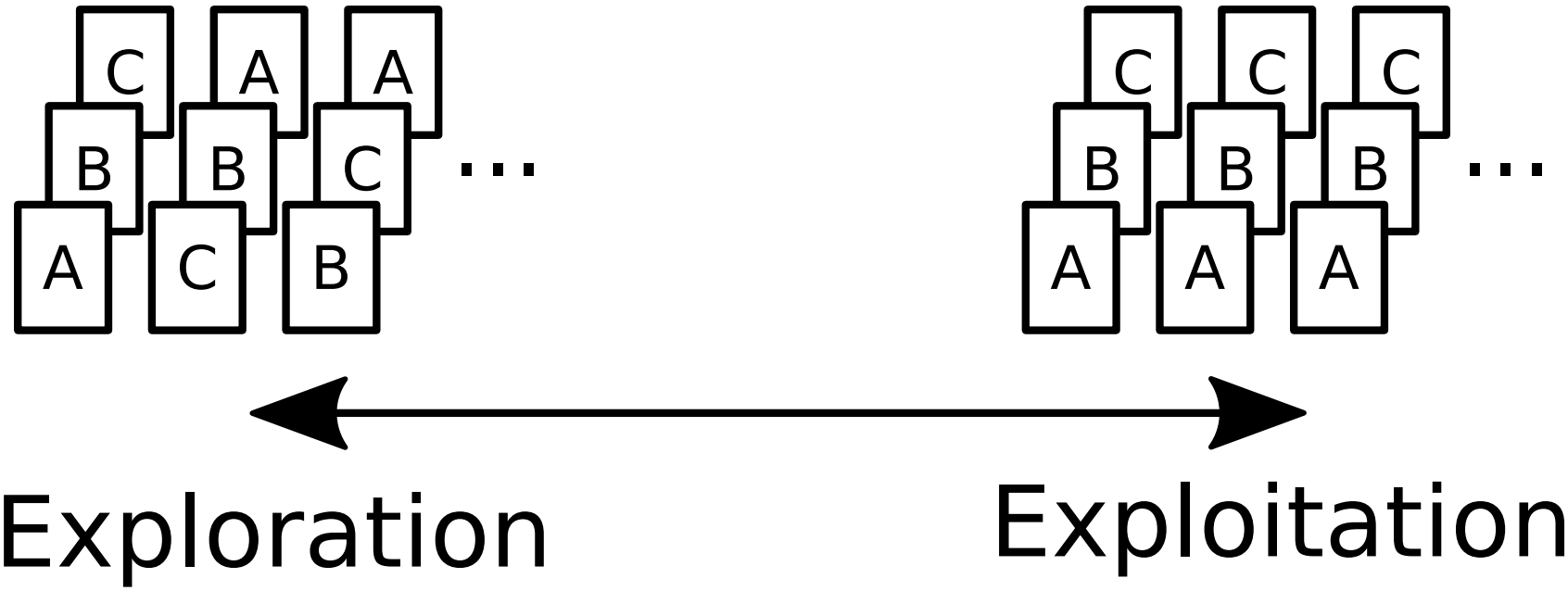




Is It Relevant to Me? (cont.)



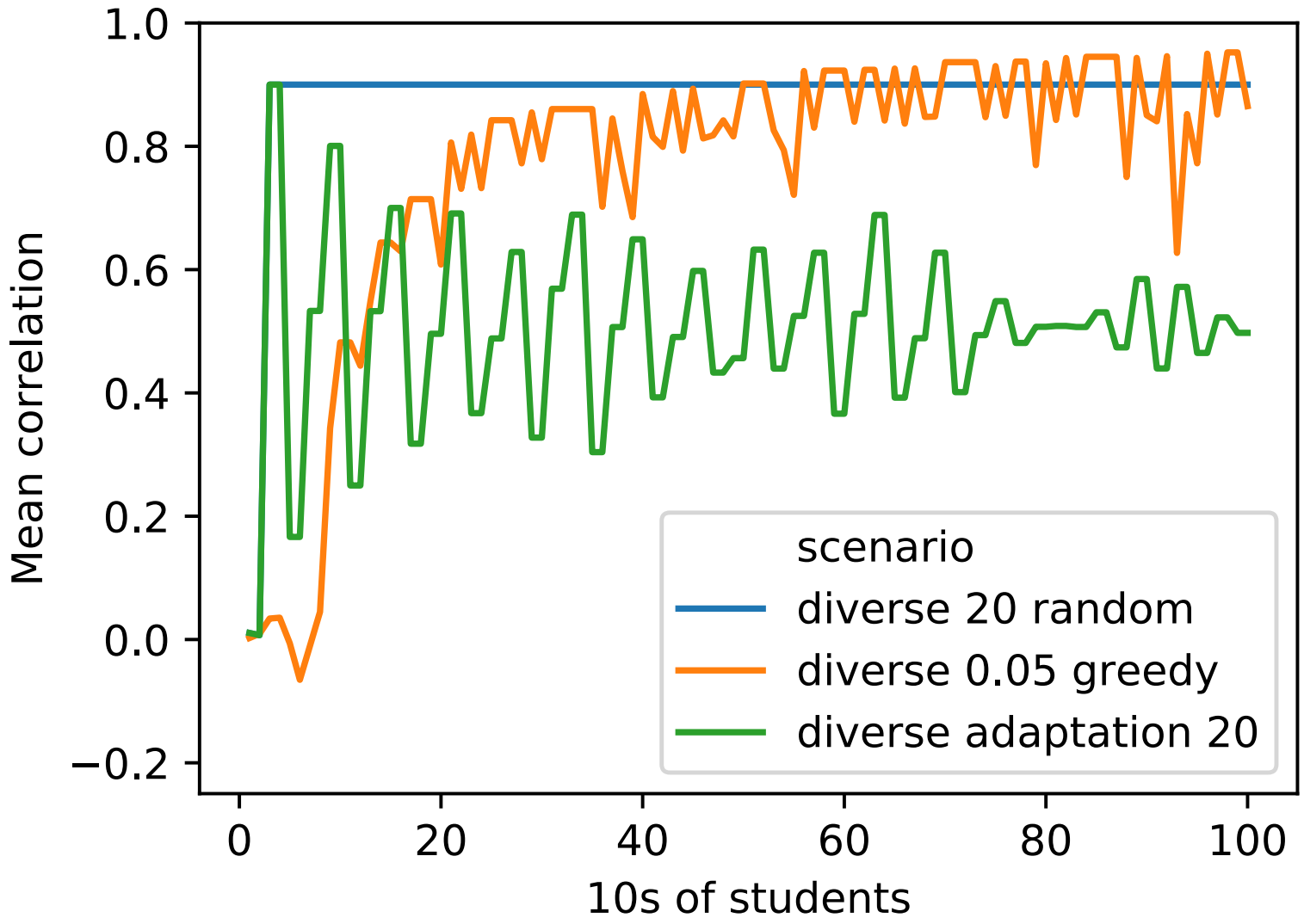
Exploration–Exploitation Tradeoff



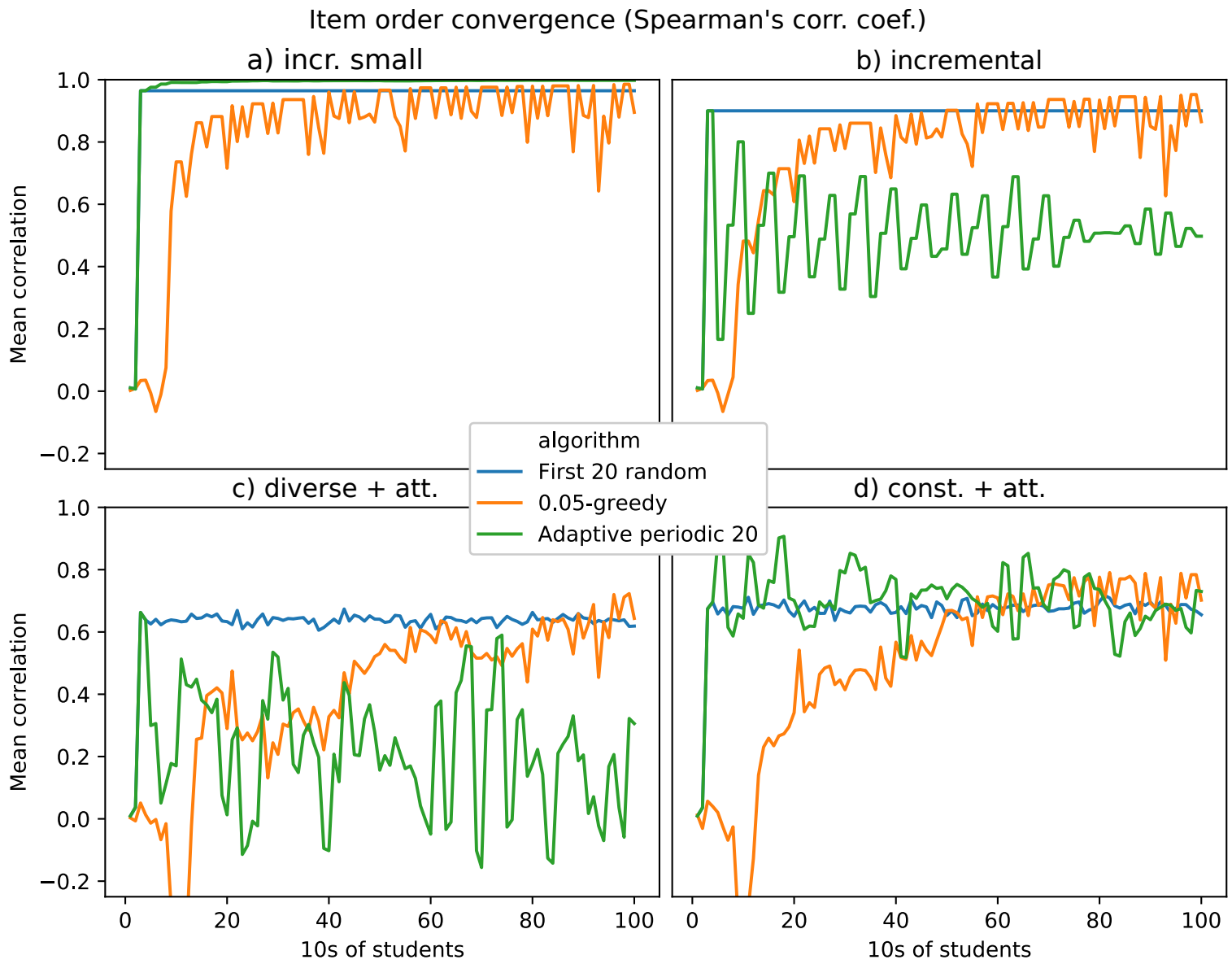
Experiments With Dynamic Ordering

- first k random
- ϵ-greedy
- adaptive periodic k

Experiments With Dynamic Ordering (cont.)



Experiments With Dynamic Ordering (cont.)



Conclusion

- item ordering bias is real
- randomization helps to reduce the bias
- know your biases before you analyze the data