Item Ordering Biases in Educational Data

Jaroslav Čechák and Radek Pelánek

Research group, Masaryk university Brno

AIED 2019

Biases in Educational Data

- implementation specific
- attrition, mastery
- **item ordering**

Why Does It Matter?

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

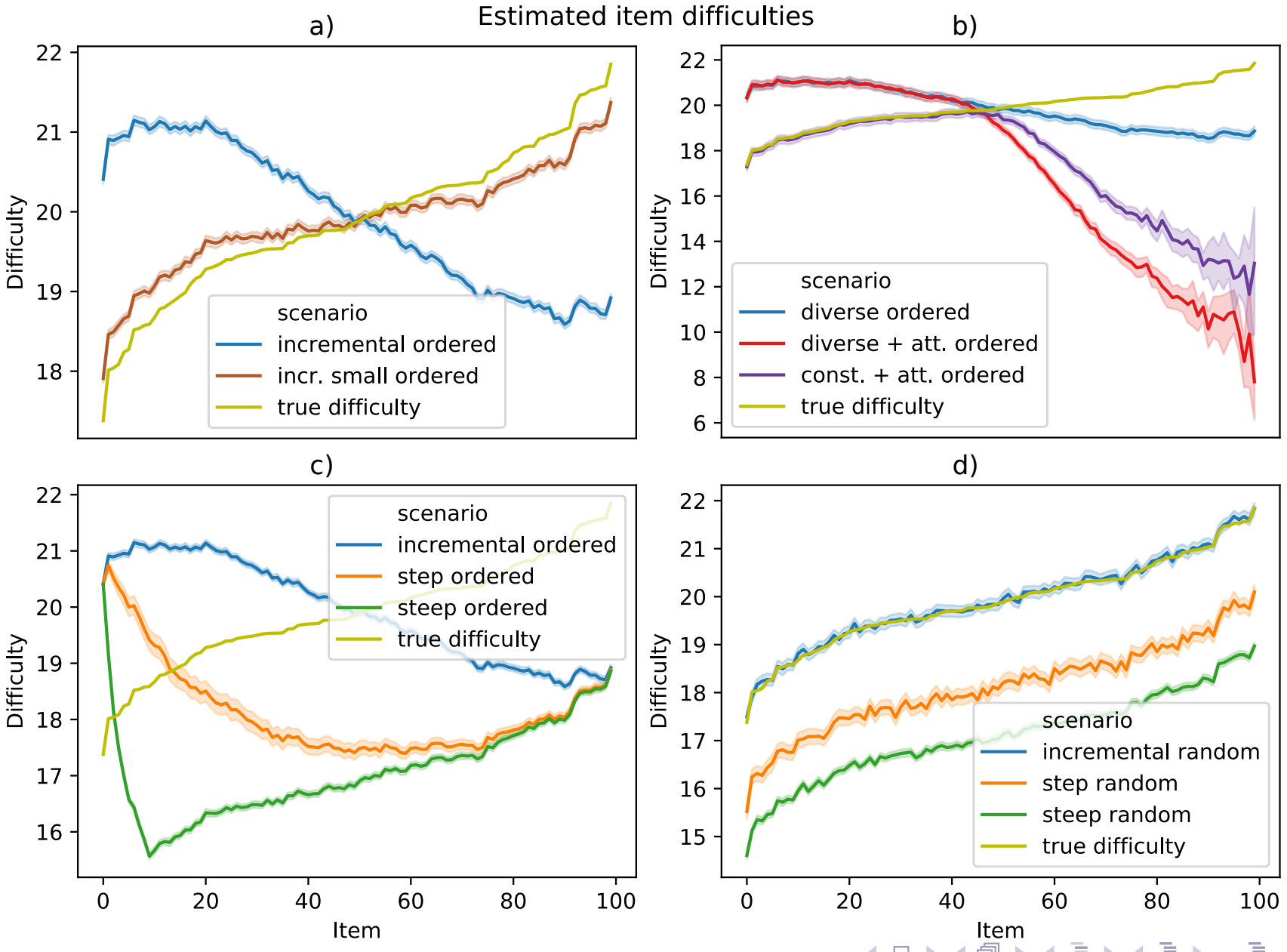
How to Explore It?

- almost impossible 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



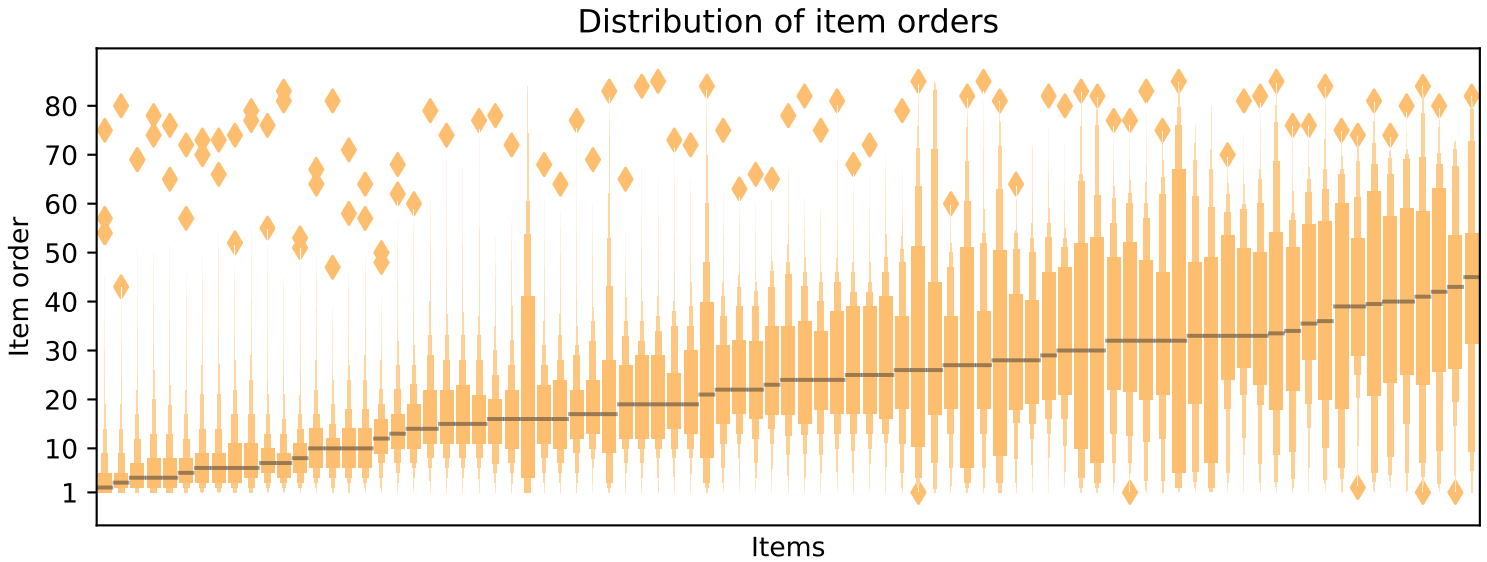
Simulation Results – Takeaway

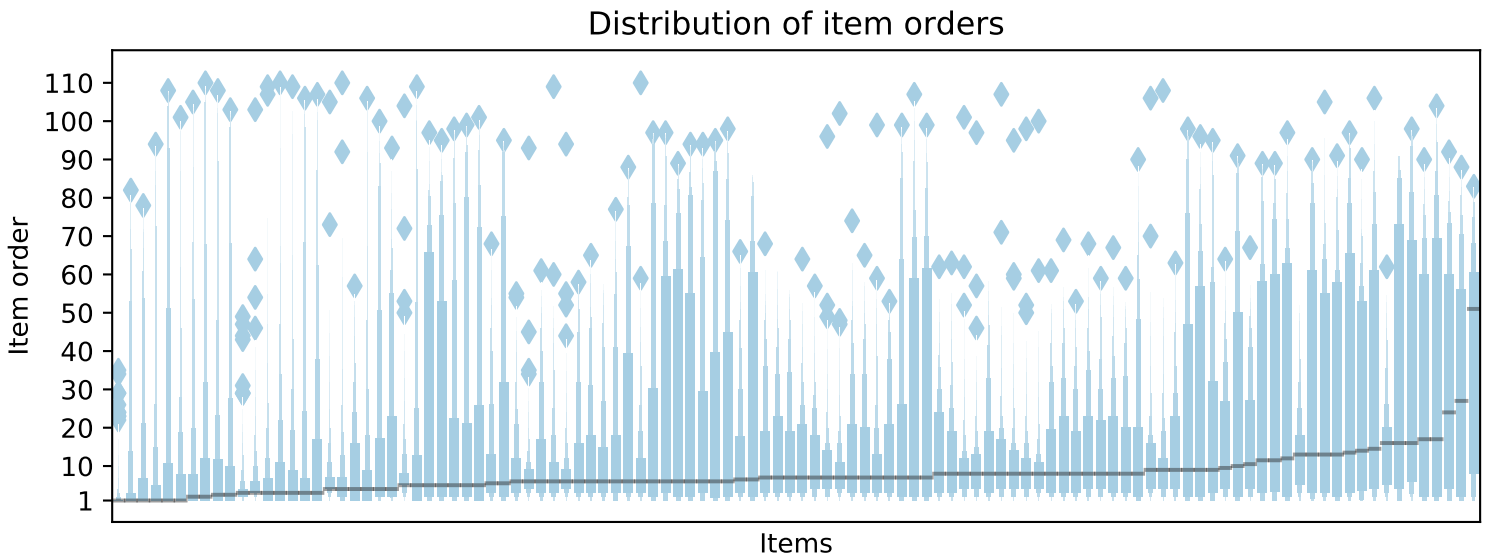
- relation between skill and difficulty is important
- randomization helps
- attrition can be a amplify the problem

Is It Relevant to Me?

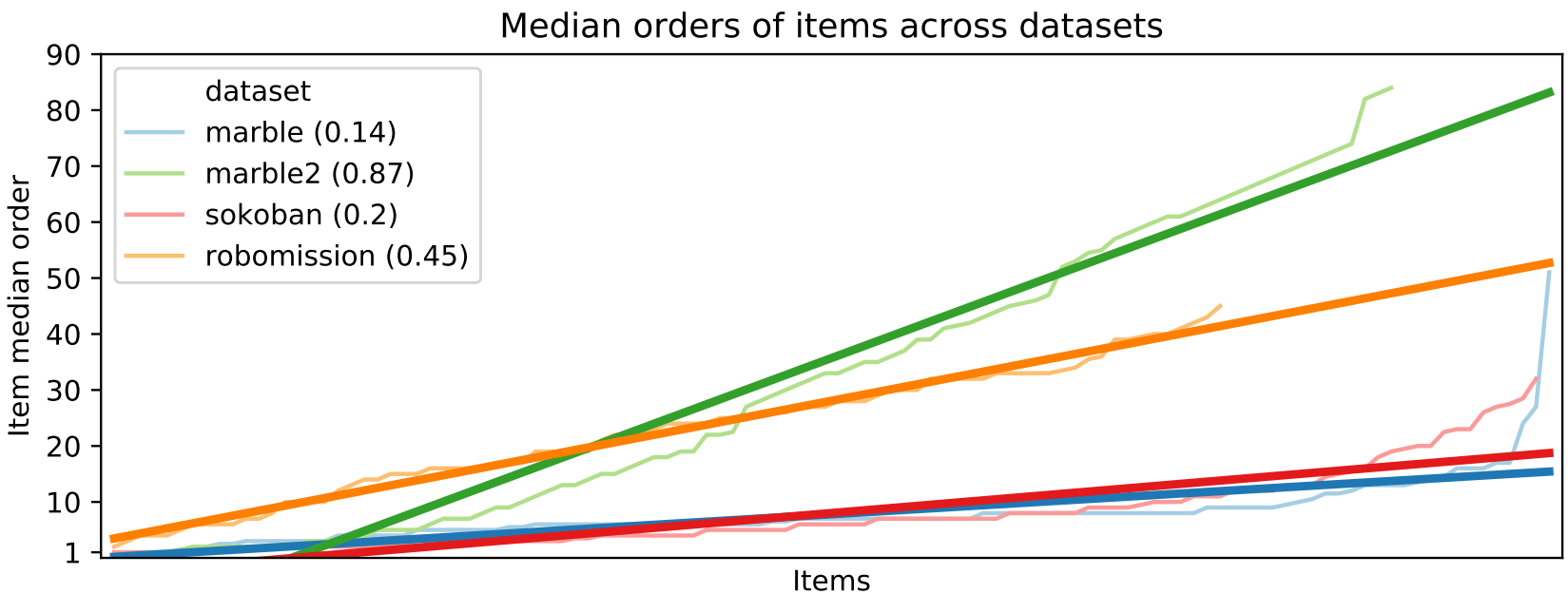
- it depends
- are student solving sequences random

Is It Relevant to Me? (cont.)





Is It Relevant to Me? (cont.)



How to Combat the Bias

- we want more randomization to reduces bias
- and student to learn efficiently

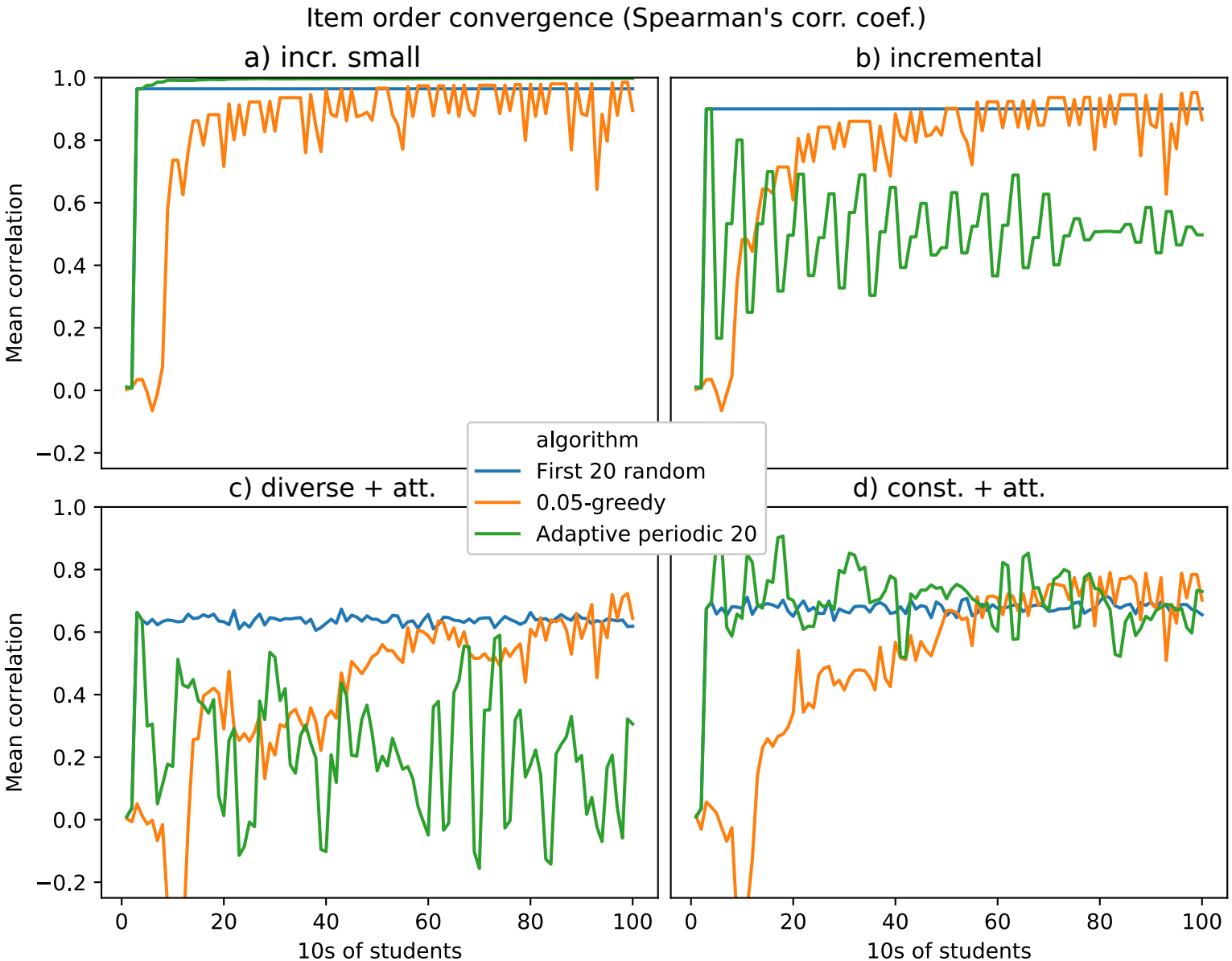
How to Combat the Bias

- we want more randomization to reduces bias (explore)
- and student to learn efficiently (exploit)
- explore-exploit tradeoff

Experiments With Dynamic Ordering

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

Experiments With Dynamic Ordering (cont.)



Conclusion and Future Work

- item ordering bias is real
- finding better ways of detecting it
- more robust methods to overcome it