Teacher vs. Algorithm

Double-blind experiment of content sequencing in mathematics

Design

Based on 3 groups

Data Collection

Teacher

Phase II

Personalized

Algorithm

Teacher

Data Collection

Phase I

Pre

Methodology

- 4th-grade (N=46), 5th-grade (N=31) students (1 Israeli school)
- Data collection December 2018-January 2019

- Learning Platform
  - International company
  - Elementary school mathematics
  - >1K interactive applets covering the curriculum
  - Each applet is scored separately
  - Multiple attempts allowed
  - Teachers can assign content to students

- Historical data
  - 15,000 Israeli students
  - ~1,000 applets
  - 564K interactions (only first attempts)
  - "Absolute difficulty" calculated by average score (tie-breaking using time spent)

Findings & Conclusions

Comparing "absolute difficulty" of recommendations

<table>
<thead>
<tr>
<th>Grade</th>
<th>Avg. (SD) Teacher</th>
<th>Avg. (SD) Algorithm</th>
<th>Paired t-test</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 (N=46)</td>
<td>2.37 (0.21)</td>
<td>1.96 (0.13)</td>
<td>9.32**</td>
<td>0.05</td>
</tr>
<tr>
<td>5 (N=31)</td>
<td>2.02 (0.44)</td>
<td>1.77 (0.18)</td>
<td>3.84**</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Comparing "absolute difficulty" and score per LO

- No difference in post-test between Teacher/Algorithm groups
- Overall, teacher’s recommendations are more difficult than algorithm’s
- When teacher’s recommendations are ordered increasingly (by “absolute difficulty”) – Teacher group performs better than Algorithm group; otherwise – no clear advantage to neither group
- Teacher may benefit from "absolute difficulty"; algorithm may benefit from content-related parameters