Discrepancies in Chapter Difficulty and Outcome

Deciphering student engagement patterns in CourseKata

Eddy Ding
Daniel Henderson
Maxim Chadaev
Siddhant Borkar
A Glimpse at Engagement & End Of Chapter Assessment Scores

P-Value = 0.028
Correlation = 0.549

Disjointedly large amount of engagement yet low EOC scores?
Low engagement, yet higher EOC scores?

EOC Score and Average Engagement for Each Chapter

Difference Between Engagement and EOC Score for Each Chapter

Low EOC Scores ≠ Difficulty
Determining the Difficulty of Chapters

Expected = Avg. Engagement Mins. Per Page + Mean of Score Engagement Differential

Relative difference Between Expected and Actual EOC. Determines the difficulty of chapters relative to all other chapters. Relative Difficulty Metric

Relative Difficulty Metric

Relative high difficulty, high # of average revisits to this chapter

Relative low difficulty, low # of average revisits to this chapter.

Revisit Submission: When a student completes the end of chapter assessment, yet still revisits the chapter and submits question submissions. More revisits suggest the student requires additional enrichment in chapter due to difficulty / lack of understanding.

Observation:

Harder difficulties (lower RDM) seemingly correlate with more average revisits to the chapter’s questions after student completion.

Disproportionately difficult chapters (chapters 2 and 3). Definite consideration for a simplification or reorganization of course materials for these chapters.
The Dynamic Difficulty Model Solution

### Statistical Analysis:
- Indicates significant performance boost in the experimental group

### Correlation Insight:
- Strong positive correlation ($r = 0.63$) between engagement and performance

### Suggestions:
- Personalized AI learning models can be successfully utilized to provide increased student performance and offset module difficulty, bringing the relative difficulty metrics closer to 0.
- The relative difficulty metric can be parameterized in an AI model to adjust difficulty of student interaction with module to ensure a more linear learning curve.

### Experimental Group:
Significant Improvement with Cohen's $d = 2.70$ (large effect)

### Control Group:
Modest Improvement with Cohen's $d = 0.44$ (medium effect)

### Table 1: Calculation of Pre-Assessment and Post-Assessment Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-Assessment</th>
<th>Post-Assessment</th>
<th>Change (Post-Pre)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>60.4 ± 7.1</td>
<td>72.8 ± 7.3</td>
<td>12.4 ± 5.3</td>
</tr>
<tr>
<td>Control</td>
<td>59.3 ± 6.5</td>
<td>64.7 ± 7.4</td>
<td>5.4 ± 3.1</td>
</tr>
</tbody>
</table>