PRACTICAL WAYS TO ACCURATELY MEASURE KNOWLEDGE AND LEARNING IN YOUR CLASSROOM

Ben O. Smith
University of Nebraska at Omaha

Jadrian J. Wooten
Virginia Tech
**What** In this workshop we will teach you to measure knowledge and learning in your classroom.

**Why** You might want to measure learning in order to: determine the effectiveness of an educational technique, measure learning for assessment, develop ways of evaluating teaching, etc.

**How** We will show you how to use the Assessment Disaggregation software for exam data, and Project Based Assessment for rubric-based instruments.
What are some ways NOT to measure learning and knowledge

- Look at exam scores
- Do a pre-and post-test and look at the difference
- Calculate average scores on rubric instruments
- Use a threshold metric: E.g. 80% of students scored >70%
What are some ways NOT to measure learning and knowledge

- Look at exam scores
  - Problem: Students guess!
  - Problem: Students have existing knowledge!
- Do a pre-and post-test and look at the difference
- Calculate average scores on rubric instruments
- Use a threshold metric: E.g. 80% of students scored >70%
What are some ways NOT to measure learning and knowledge

- Look at exam scores
  - Problem: Students guess!
  - Problem: Students have existing knowledge!
- Do a pre-and post-test and look at the difference
  - Problem: Students guess and they guess at different rates!
- Calculate average scores on rubric instruments
- Use a threshold metric: E.g. 80% of students scored >70%
THE PROBLEM WITH POST-PRE

pre-test guessing

\[ \mu \]

post-test guessing

\[ \gamma + \mu \]
What are some ways NOT to measure learning and knowledge

- Look at exam scores
  - Problem: Students guess!
  - Problem: Students have existing knowledge!
- Do a pre-and post-test and look at the difference
  - Problem: Students guess and they guess at different rates!
- Calculate average scores on rubric instruments

- Use a threshold metric: E.g. 80% of students scored >70%
What are some ways NOT to measure learning and knowledge

- Look at exam scores
  - Problem: Students guess!
  - Problem: Students have existing knowledge!
- Do a pre-and post-test and look at the difference
  - Problem: Students guess and they guess at different rates!
- Calculate average scores on rubric instruments
  - Problem: Rubric rows are ordinal and the top box is censored
  - Problem: Rubric weights impact the overall scores
- Use a threshold metric: E.g. 80% of students scored >70%
What are some ways NOT to measure learning and knowledge

- Look at exam scores
  - Problem: Students guess!
  - Problem: Students have existing knowledge!
- Do a pre-and post-test and look at the difference
  - Problem: Students guess and they guess at different rates!
- Calculate average scores on rubric instruments
  - Problem: Rubric rows are ordinal and the top box is censored
  - Problem: Rubric weights impact the overall scores
- Use a threshold metric: E.g. 80% of students scored >70%
  - Problem: Threshold is arbitrary
  - Problem: Threshold represents an information reduction
Problems with Threshold Measures

- Measure of percent of students above a score
- Not centered on truth
- Throws away data
PROBLEMS WITH THRESHOLD MEASURES
Today we’ll show you that you can properly measure learning and knowledge, and it does not require that much effort.
USING EXAMS TO PROPERLY MEASURE LEARNING
Students who *learned* the material ($\gamma$)
Students who already *knew* the material ($\mu$)
Students who guessed

---

Exam Answering Machine

$1$

$0$
1. Pre-test: Because otherwise, we can’t distinguish between students who knew the material before the class and those who learned it in your class.

2. Post-test: Because need to know how much the students know at the end of the class.

3. Mapping: Because we need to match the pre-test questions with the post-test questions.

4. The probability of guessing correctly: Because some students may have answered correctly by chance!
Assessment Disaggregation

Required Files
To perform this disaggregation, you must specify a pretest and posttest with some set of matched questions. Standard Scantron, Akindi, Quick Key, ZipGrade, Moodle, Google Form quizzes, Canvas, and some Blackboard formatted files can be specified. For more information about the exam file formats, please see the help menu for the documentation.

Optional Files
The assessment map is optional but recommended. This file's columns are "q" (assessment question number in output), "exam1" (pretest question number), "exam2" (posttest question number), and "options" (number of answer options for the given question). Student list limits the students used for the analysis using a singular column: id.

Pretest
-  Posttest
-  Matched Students

I need more data!
To perform an analysis, you need to load the pretest and posttest. If you don't specify an assessment map file, I will assume the questions are in the same order and all questions have four options. If you don't specify the student list file, I will use all students who appear in both the pretest and posttest.
1. Assessment Disaggregation:
   https://www.assessmentdisaggregation.org

2. Some Demo Files:
DISAGGREGATION DEMO
USING RUBRIC-BASED INSTRUMENTS TO MEASURE KNOWLEDGE
### Example Rubric

<table>
<thead>
<tr>
<th>Did not meet any of the requirements</th>
<th>The chosen topic was on-topic for the course, but the literature was lacking</th>
<th>The literature review was complete but contained errors</th>
<th>Full credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 0$</td>
<td>$k = 1$</td>
<td>$k = 2$</td>
<td>$k = 3$</td>
</tr>
</tbody>
</table>
**Needed Ingredients**

1. The $k_{ij}$/score value: The box they landed in from lowest (0) to highest ($b_j$)
2. Bound/$b_j$: The maximum $k_{ij}$ value possible on a given rubric row
3. Identifier for the student
4. Identifier for the rubric row
5. Subset (optional): A list of students to treat as a separate group

Items (1) - (4) are simply in the same CSV file. Item (5) is a separate file.
## Project Based Assessment

### Rubric Information

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Average Logistic</th>
<th>Average Marginal Logistic</th>
<th>Average Discrete Marginal Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.681</td>
<td>0.067</td>
<td>-0.068</td>
<td>-0.131</td>
</tr>
<tr>
<td>2</td>
<td>0.587</td>
<td>0.286</td>
<td>0.099</td>
<td>0.087</td>
</tr>
<tr>
<td>3</td>
<td>0.811</td>
<td>0.325</td>
<td>0.149</td>
<td>0.127</td>
</tr>
<tr>
<td>4</td>
<td>-0.153</td>
<td>0.180</td>
<td>-0.018</td>
<td>-0.019</td>
</tr>
<tr>
<td>5</td>
<td>0.612</td>
<td>0.290</td>
<td>0.105</td>
<td>0.092</td>
</tr>
<tr>
<td>6</td>
<td>0.817</td>
<td>0.326</td>
<td>0.150</td>
<td>0.128</td>
</tr>
<tr>
<td>7</td>
<td>0.918</td>
<td>0.345</td>
<td>0.174</td>
<td>0.147</td>
</tr>
<tr>
<td>8</td>
<td>-0.472</td>
<td>0.145</td>
<td>-0.045</td>
<td>-0.053</td>
</tr>
</tbody>
</table>

### Model Fit

### Student KDE Estimates
Stuff You Need

1. Project Based Assessment: https://projectassessment.app
PROJECT BASED ASSESSMENT DEMO
Both software packages have documentation sites:
- https://docs.assessmentdisaggregation.org
- https://docs.projectassessment.app

The methods are based on statistics:
- Assessment Disaggregation
- Project Based Assessment - https://bit.ly/a-rubric-d