Using Data to Inform Standard-Setting Recommendations

Maryland Integrated Science Assessment, Grades 5 & 8

Jennifer Judkins, MSDE   |   Matt Gaertner, WestEd
PART 1

Background
Background

2013
The Maryland State Board of Education adopts the Next Generation Science Standards

2016-2017
The Maryland Integrated Science Assessment (MISA) replaces the Maryland School Assessment in science
Background

Spring 2018
The MISA is administered for the first time operationally

Summer 2018
WestEd and MSDE conduct standard-setting meetings for grade 5 and grade 8 assessments
Background

Four MISA Performance Levels

<table>
<thead>
<tr>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partially Met Expectations</td>
<td>Approaching Expectations</td>
<td>Met Expectations</td>
<td>Exceeded Expectations</td>
</tr>
</tbody>
</table>
Background

### Three MISA Cut Scores

<table>
<thead>
<tr>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partially Met Expectations</td>
<td>Approaching Expectations</td>
<td>Met Expectations</td>
<td>Exceeded Expectations</td>
</tr>
</tbody>
</table>
PART 2

Meeting Preparation
Meeting Preparation

Key Considerations for Using Data During Standard Setting

• **What** data should be shared with panelists?

• **When** should those data be shared?

• **How** should those data be shared?
Meeting Preparation

What data?

► When convergence is the goal, panelists need to understand how their judgments compare to their colleagues’ judgments

► Panelists have a right to know the impact of their recommendations

► Panelists are science educators, but NGSS assessments are new, so student performance is difficult to forecast
Sharing data with panelists: What, when, and how?

Meeting Preparation

When?

- Standard setting for criterion-referenced tests should begin with content – students’ knowledge and skills relative to the standards

- Panelists judgments converge incrementally through deliberation – informed but not dictated by data

- Impact data (e.g., projected proficiency rates) should be a reality check, not a starting point
Meeting Preparation

How? (… and how much?)

- Without explanation and context, summary statistics (e.g., p-values) are ambiguous at best

- Standard-setting data (just like items on a test) should produce the intended interpretations

- With new standards, new assessments, and three cut scores, panelists have a lot to consider. Data overload is a real concern.

Sharing data with panelists: What, when, and how?
Meeting Preparation

1. Student Data
2. A Bit of Code
3. Item/Pra Software
PART 3

Standard Setting Meeting
Standard Setting Meeting

Three Rounds of Judgment

Round 1 Data
Judgments from your table and your grade

Round 2 Data
Judgments + item p-values

Round 3 Data
Judgments + p-values + impact data (overall and by demographic groups)
Standard Setting Meeting

Vertical Articulation

After all the judgments were complete, grade 5 and grade 8 panelists reconvened and reviewed the cut scores together, with two fundamental questions in mind...
Standard Setting Meeting

Vertical Articulation

1. From the perspectives of students, parents, educators, and the general public, do these cut scores make sense?

2. If not, how should the cut scores be adjusted?
Standard Setting Meeting

*Vertical Articulation*

The full group considered three relevant benchmarks in addition to MISA impact data:

- Impact data from similar states with similar NGSS-aligned assessments
- NAEP Science impact data
- PARCC Mathematics and English / Language Arts impact data
Thank you!

Jennifer Judkins
Assistant State Superintendent
Division of Assessment, Accountability, and Information Technology
Maryland State Department of Education

Matthew Gaertner
Director of Research
WestEd