

# Science Assessment Webinar

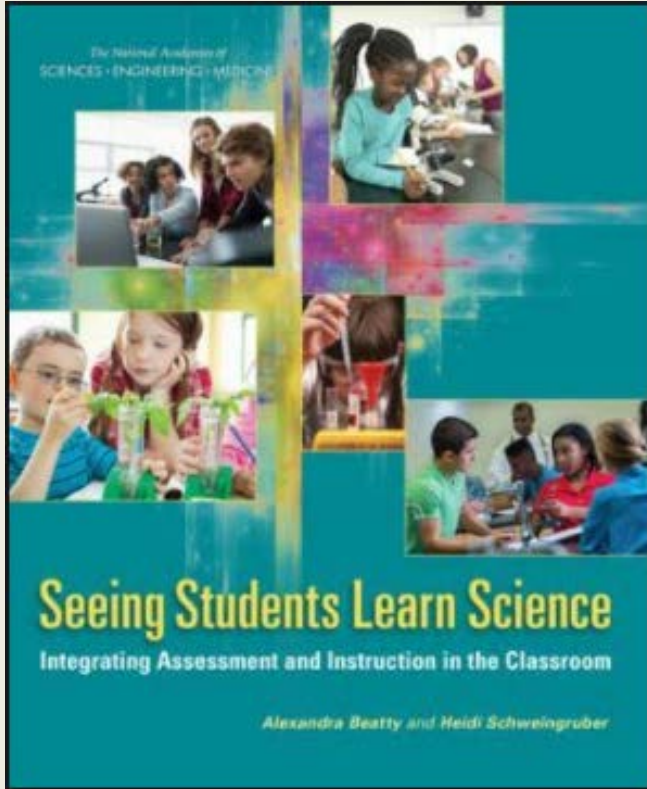
## WHAT a YEAR!



**TJ SMOLEK**  
**AUGUST 29, 2017**



# An Assessment System for Science



“The idea of an assessment system begins with a commonsense point: no one assessment – or assessment occasion – can meet all the needs for information about what students know and can do in science” (p.21 NASEM, 2017).

“Large-scale assessments, particularly the yearly tests used by districts and states, play a key role in shaping both expectations for student learning and public discussion and perceptions of science education. Therefore, it is critical that these test be adapted along with instruction” (p. 24, NASEM, 2017).

National Academies of Sciences, Engineering, and Medicine. 2017. *Seeing Students Learn Science: Integrating Assessment and Instruction in the classroom*. Washington, DC: The National Academies Press.

# Science Assessment System Goals

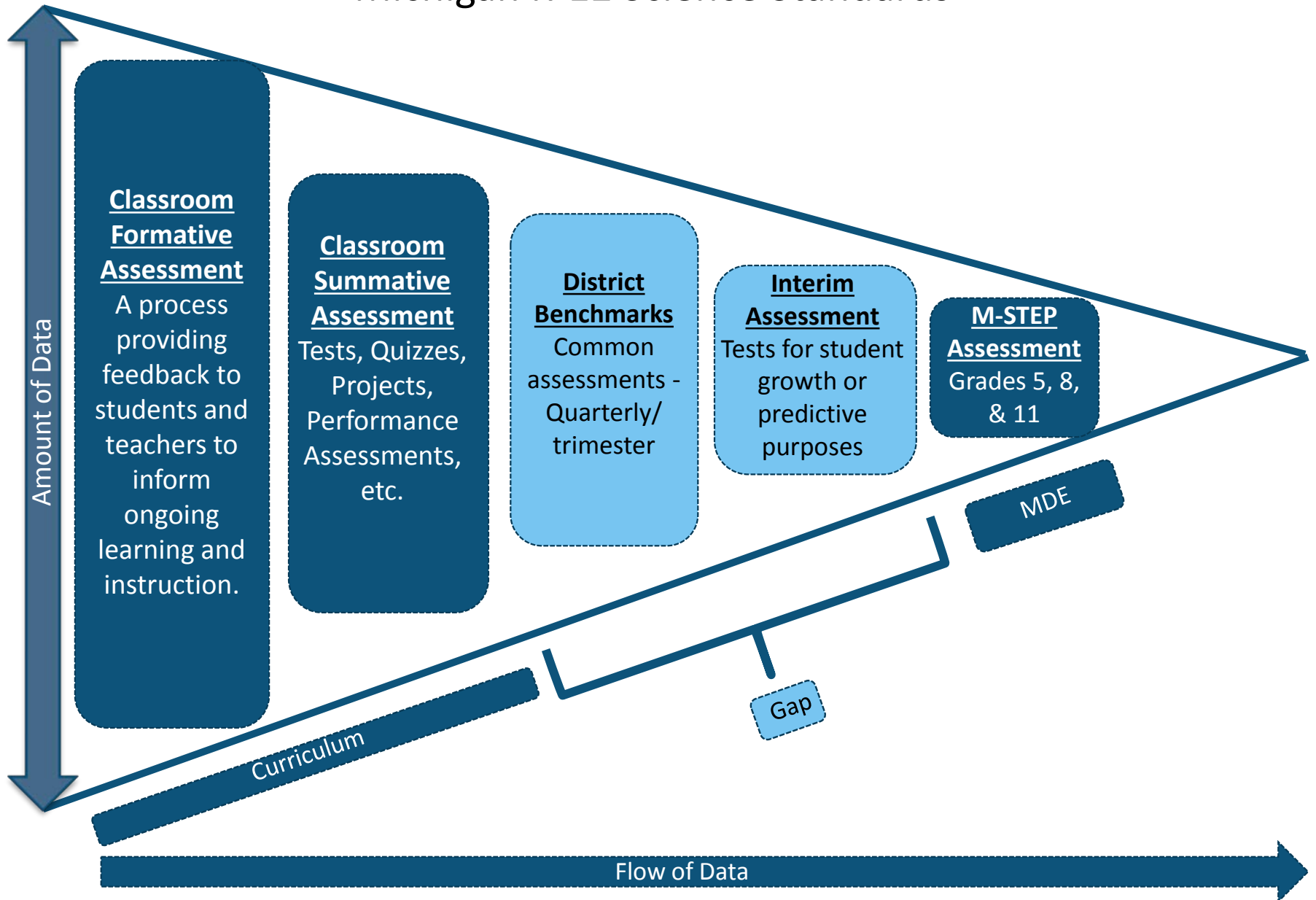


- Science assessments in Michigan must be a coherent system of assessment to support both classroom learning and policy/monitoring functions.
- Michigan monitoring (accountability) science assessments must move beyond traditional forms; testing as usual will NOT suffice.
- Opportunity to learn science is an essential system component.
- Classroom science teaching and assessment come first.

NRC, 2014

“Changing large-scale accountability tests may be the most challenging piece of the puzzle, but teachers can proceed even while system-wide changes are evolving”  
(p.22, NASEM, 2017).

# Vision for Balanced Assessment System for Michigan K-12 Science Standards



# Evidence Centered Design: Michigan Science Assessment Claims



**Equity Claim:** Non-dominant and dominant groups of students have the opportunity to demonstrate grade-band proficiency through the use of engineering, local contexts, and relevant phenomena.

**Scientific Literacy Claim:** Students demonstrate grade-band proficiency in using the three dimensions to critically evaluate scientific and technological information in order to design solutions to problems and investigate phenomena.

## **Student Level Claim:**

Student has demonstrated grade-band proficiency in:

**Life Science,  
Earth Science, &  
Physical Science**

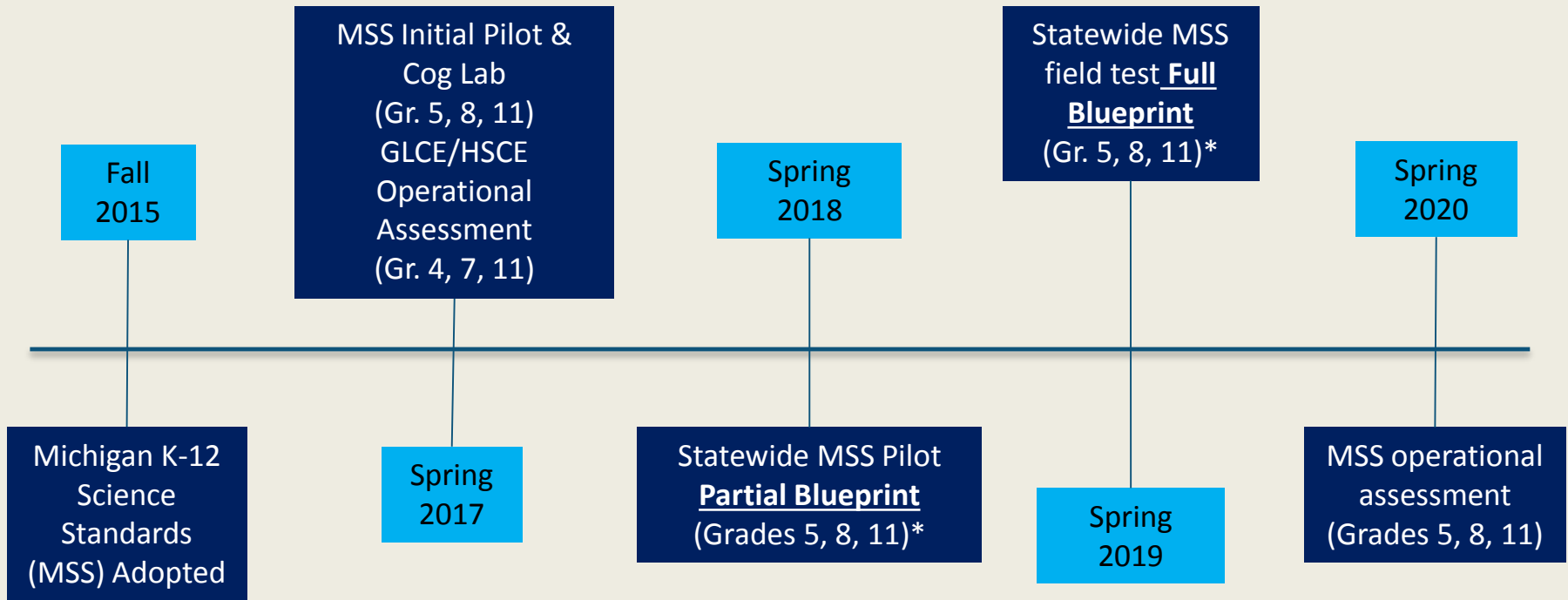
Topic Bundles using all dimensions represented in the standards.



## **District/State Level Claim:**

Students have demonstrated grade-band proficiency to explain the presented phenomenon (local or global) and design solutions to problems using all dimensions represented in the given **topic bundle**.

# Implementation Timeline



\*Accountability is based on student participation in pilot / field test in 2018 and 2019.

Operational science assessment of GLCEs and HSCEs will NO LONGER be conducted.

# 2018 – 2019 Pilot & Field Test



## 2018 – Pilot Test

- 2 Forms per grade (5, 8, 11)
- Partial Test Map
- 3 Item Clusters per form
  - 1-Physical Science
  - 1- Earth Science
  - 1- Life Science

## 2019 – Field Test

- 2 Forms per grade (5, 8, 11)
- Full Test Map
- 6 Item Clusters per form
  - 2-Physical Science
  - 2- Earth Science
  - 2- Life Science

Develop and include student supports:  
Paper-pencil form, Text-to-Speech, Accommodated form, Braille form, Translations

The pilot and field tests will provide important information regarding actual testing times so OEAA can make the necessary adjustments

# 2020 Proposed Operational Test



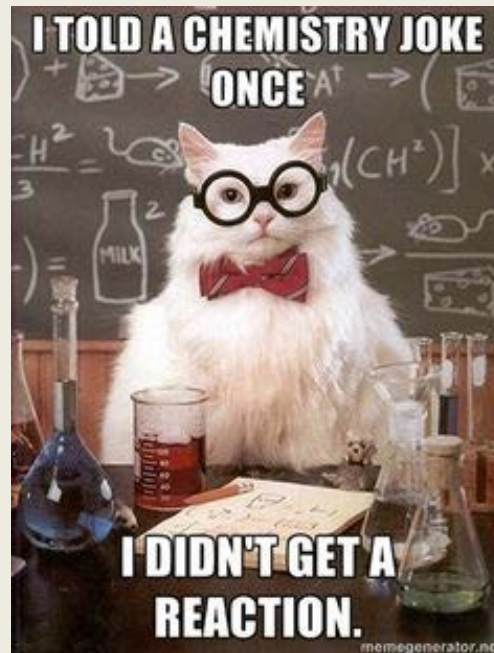
- 3 Forms per grade (5, 8, 11)
  - Full Test Map
  - 7-8 Item Clusters per form
    - 2-Physical Science
    - 2- Earth Science
    - 2- Life Science
    - 1-2 – Field Test Item Cluster(s)
- } Operational



# 2017 Pilot Test



CONDUCTED APRIL 11-MAY 26, 2017



# 2017 Pilot Test



- 21,469 participants
  - Grade 5 – 6,732
  - Grade 8 – 9,331
  - **Grade 11 – 5,406**
- Somewhat representative of demographic diversity
- Somewhat representative of geographic diversity
- Student surveys accompanied pilot tests.

# Gender & Ethnicity Demographics



## Science Pilot (Grades 5, 8, 11)

Category	N Count	%
Male	10,601	49
Female	10,868	51
Hispanic/ Latino	1,142	5 (-2)
American Indian or Alaska Native	98	>1
Asian	609	3
Black or African American	1870	9 (-8)
Native Hawaiian or Other Pacific Islander	15	>1
White	16,892	79 (+12)
Two or More Races	815	4

## Michigan Students (Grades 5, 8, 11)

Category	N Count	%
Male	176,074	51
Female	170,562	49
Hispanic/ Latino	25,121	7
American Indian or Alaska Native	2,261	>1
Asian	11,418	3
Black or African American	58,947	17
Native Hawaiian or Other Pacific Islander	324	>1
White	232,558	67
Two or More Races	11,288	4

# More Demographics



## Science Pilot (Grades 5, 8, 11)

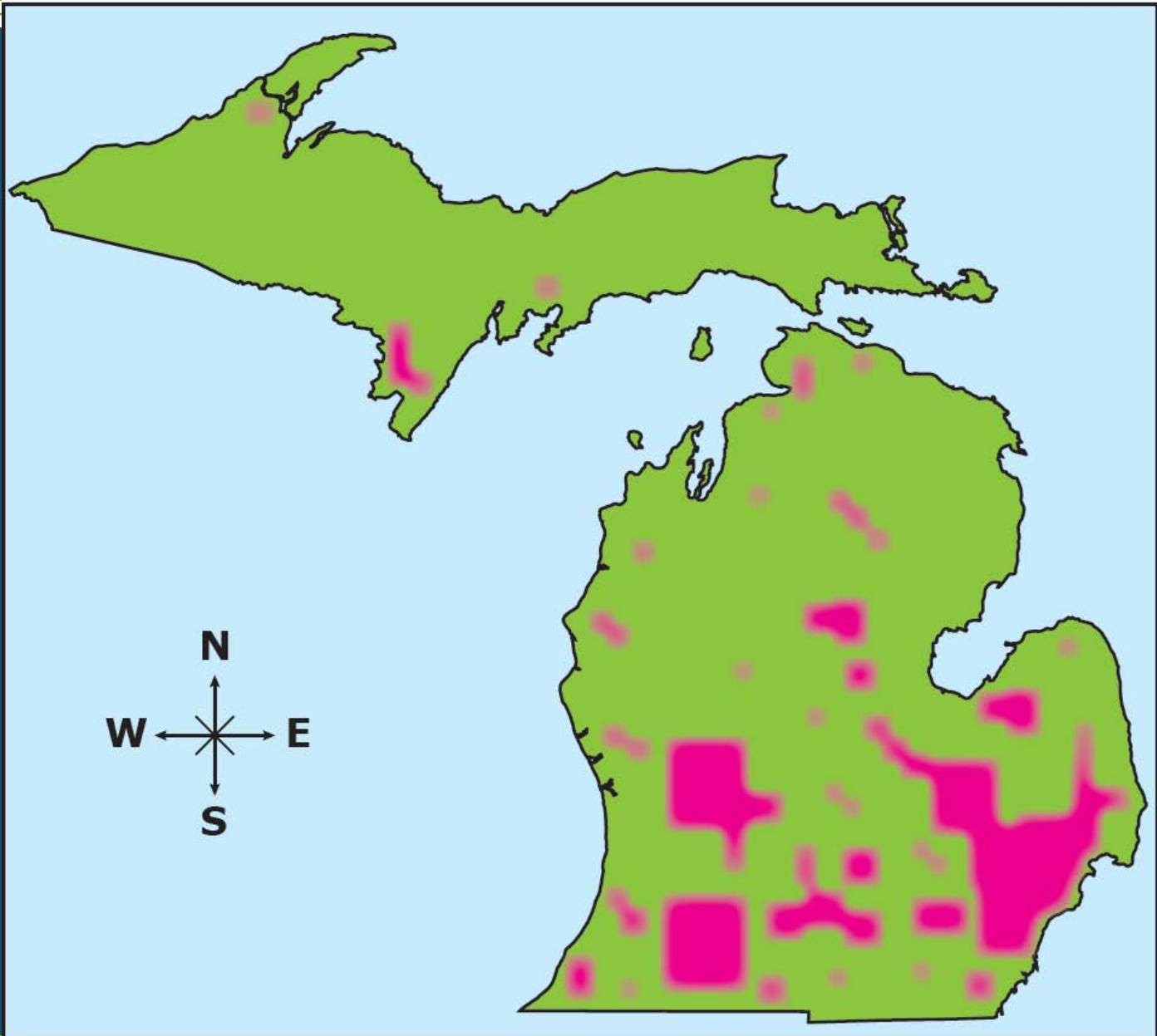
Category	N Count	%
Economically Disadvantaged	6,778	31 (-12)
English Learner	342	2 (-3)
Students with Disabilities	1,796	8 (-4)

## Michigan Students (Grads 5, 8, 11)

Category	N Count	%
Economically Disadvantaged	148,361	43
English Learner	17,776	5
Students with Disabilities	42,829	12



# Pilot Test Geographical Representation



# Constructed Response Scoring



- 7 constructed response items – some with multiple parts

Experts from  
MDE create  
anchor and  
training sets

Michigan  
educators hand-  
score items  
(August 2017)

Gain feedback  
from scorers  
regarding  
revisions of CR  
items and rubrics

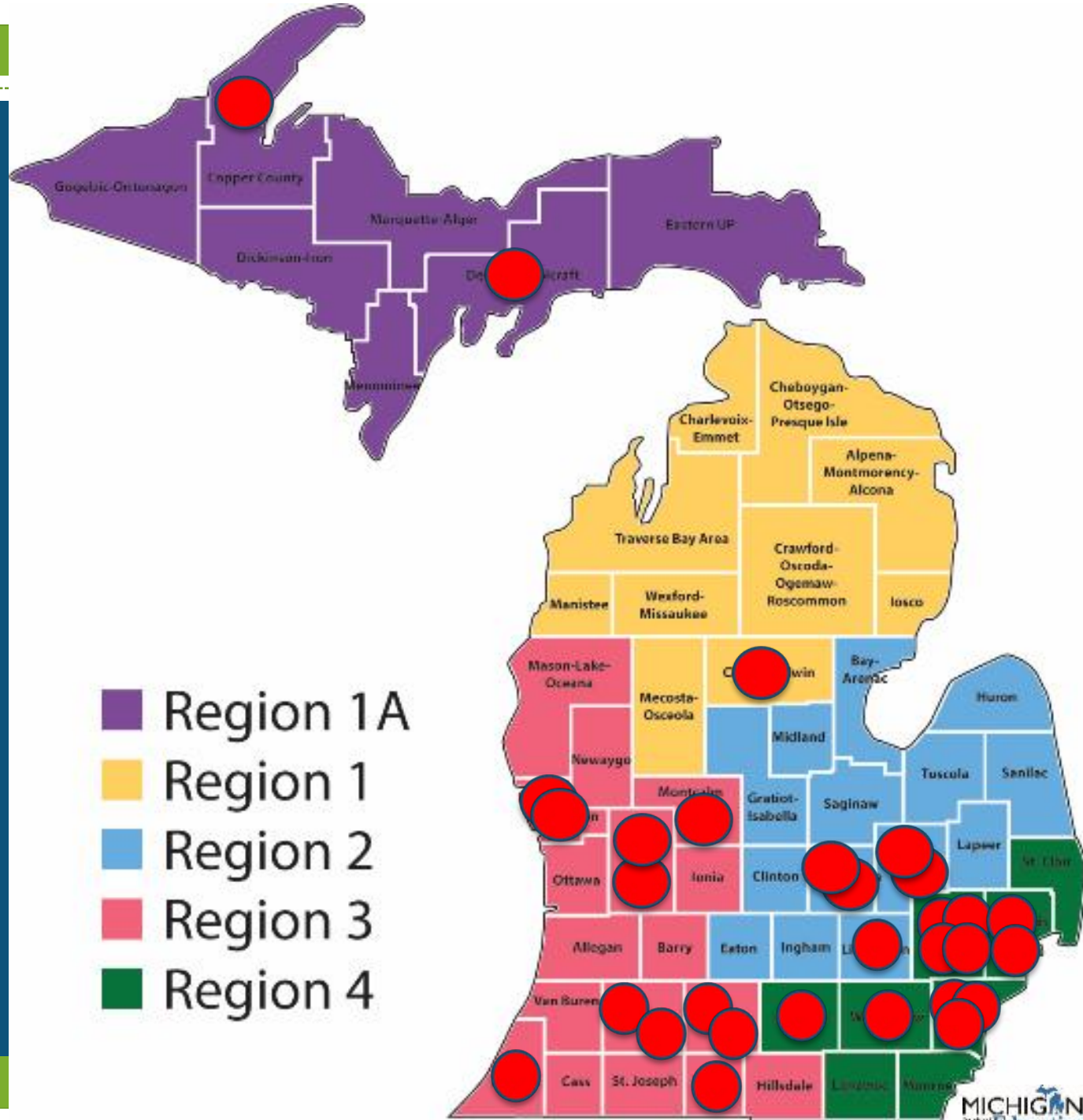
# Cognitive Labs



- Somewhat representative of geographic diversity
- No information regarding demographic diversity is available at this time as UIC numbers were not collected from students to protect their privacy.
  - Grade 5 - 20 students
  - Grade 8 – 22 students
  - Grade 11 – 29 students



# Cognitive Lab Geographic Distribution





# Field Note Themes



**Asked for a range of students – got the best and brightest**

**Students liked that they could “figure out” the answers**

**Multiple click to enlarge windows were frustrating**

**Second stimulus wasn't always noticed by the student**

**Less advanced students drew on prior knowledge more than the data presented**

**Stimulus wasn't always necessary to answer question**

**Students LOVED the graphics**

**“I liked the colorful pictures and real world topics. It was like it mattered to me.”** (Grade 11 Cognitive Lab participant, May, 2017)

# Limitations and Opportunities for Learning



- Sample size restricts generalization capabilities
- Fine-grain analysis will be conducted
- Eliciting 3-D thinking is not easy
- Mini clusters may have to be considered
- Item clusters will pose challenges to typical standard-setting process

# Release Annotated Item Clusters – Fall 2017



**Gr 5 Science**  
Question 1

**Lights Out! (Part 1)**

Students are learning about eyes in science class. During their class discussion, a power outage occurs and the lights go out in the classroom. While the teacher looks for a flashlight, one student exclaims, "I can't see anything!"

The teacher turns on a flashlight and points it across the classroom to a plant on a table. The teacher says, "This makes me wonder how we are able to see the plant."

Which statement **best** describes how the students are able to see the plant?

- Once the plant produces its own light, the students can observe the plant.
- Once the plant absorbs all the light from the flashlight, the students can observe the plant.
- The light from the flashlight is reflected toward the students' eyes and then back to the plant.
- The light from the flashlight is reflected off of the plant and then enters the students' eyes.

More Text Below

**Gr 8 Science**  
Question 1

**How do Hand Warmers Work? (Part 1)**

Two students are outside in the cold, waiting for a bus. One of the students has a package of hand warmers and offers to share them with the other student. The student opens the package and they each put a hand warmer bag in one of their gloves.

After a few minutes, the students notice that the hand warmer bags start to feel warm. The students want to know how hand warmer bags get warm. They decide to ask their science teacher if they

This question has two parts.

**Part A:** Use the data table to complete the following statement.

The students can tell that a chemical reaction involving iron \_\_\_\_\_ because a new substance \_\_\_\_\_ form overnight.

**Part B:** Choose one set of properties that **best** supports the completed statement in Part A.

- density and color
- color and volume
- volume and texture
- texture and mass
- mass and density

**Gr 11 Science**  
Question 1

**Atmospheric Changes over Time (Part 1)**

The gases that make up Earth's atmosphere have changed over time. Scientists measure the levels of carbon dioxide ( $\text{CO}_2$ ) in Earth's atmosphere. The model shown represents a simplified version of the carbon cycle. Some of the locations where carbon is stored are identified and processes that move carbon from one location to another are indicated.

Scientists can use carbon cycle models to help make predictions about the amounts of carbon in different locations. Use this carbon cycle model to identify **all** the processes that would decrease  $\text{CO}_2$  in the atmosphere if the rate of that process were to increase.

- photosynthesis
- cellular respiration
- fossil fuel combustion
- diffusion into the ocean
- diffusion into the atmosphere

# Future Tools and Research



- Item Cluster Alignment Tools
- Item Cluster Evaluation Tool
- Analyze Cognitive Lab data and Pilot Test data for gaps
- Continue transparency about the need for continuous improvement with the assessments

# Item Cluster Development as Professional Learning



Year 1 - 2016	Year 2 - 2017
5 weeks of ICD	4 weeks of ICD
46 Science Educators	28 returning Science Educators
	46 new Science Educators
Lansing, MI	Lansing, MI and Marquette, MI

“Participating provided me with a better understanding of the student assessment trajectory - how formative and summative assessment in the classroom can support large level state assessments. It also helps me contribute to conversations about NGSS, and three-dimensional assessment” (Survey question response, Cohort 1 Participant, 2017).

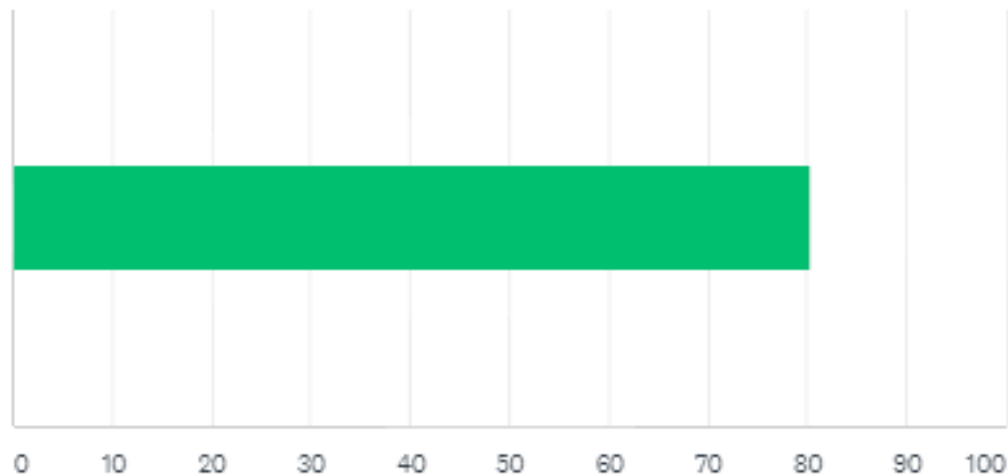


# Item Cluster Development as Professional Learning



Please indicate the extent to which your experience with Item Cluster Development during the Summer of 2016 impacted your professional work over the course of the 2016-2017 school year.

Answered: 17 Skipped: 0



# Item Cluster Development as Professional Learning



**Q7** Please indicate the degree to which facets of your professional work were impacted by your experiences in the 2016 Item Cluster Development.

