Modeling Student Learning of the Structure of Matter and Scientific Argumentation: Implications for Aligning Assessments with the Next Generation Science Standards

Linda Morell, Tina Chiu, Mark Wilson (UCB); Jonathan Osborne & Andrew Wild (Stanford)

April 28, 2015

BEAR Seminar
OUTLINE

• The Project & Context
• Structure of Matter (SOM)
• Argumentation
• Relationship between SOM and Argumentation
• Questions and Comments
LEARNING PROGRESSIONS IN MIDDLE SCHOOL SCIENCE INSTRUCTION AND ASSESSMENT (LPS)

• Assessment project in Science Education
• Funded by Institute of Education Sciences (IES)
• Collaboration of four institutions
• Based on previous research
RESEARCH QUESTIONS

1. What is the nature of the learning progression in the content domain of Structure of Matter?

2. What is the nature of the learning progression in students’ ability to argue from evidence in the domain of Structure of Matter and also, generally?

3. What is the inter-relationship between students’ ability to argue from evidence and their domain specific knowledge? In particular to what extent do the two co-vary?
CONTEXT OF THE RESEARCH

• LPS proposed in 2008
• Next Generation Science Standards (NGSS, 2013)
• Developing Assessments for the Next Generation Science Standards (2014)
Core Concepts: Physical Science

- Chemical Reactions
- Forces and Motion
- Structure of Matter
- Types of Interactions
- Stability and instability in physical systems
- Definitions of energy
- Conservation of energy & energy transfer
- Relationship between energy and forces
- Energy in chemical processes and everyday life
- Wave properties
- Electromagnetic radiation
- Information technologies & instrumentation

Practices:

- Asking questions & defining problems
- Developing and using models
- Planning & carrying out investigations
- Analyzing & interpreting data
- Using math & computational thinking
- Constructing explanations & designing solutions
- Engaging in argument from evidence
- Obtaining, evaluating, and communicating information

Crosscutting Concepts:

- Patterns
- Cause and effect
- Scale, proportion, and quantity
- Systems & system models
- Energy & matter
- Structure & function
- Stability & change
NEXT GENERATION SCIENCE STANDARDS

Core Concepts: Physical Science

• Chemical Reactions
• Forces and Motion
• **Structure of Matter**
  • Types of Interactions
  • Stability and instability in physical systems
  • Definitions of energy
  • Conservation of energy & energy transfer
  • Relationship between energy and forces
  • Energy in chemical processes and everyday life
• Wave properties
• Electromagnetic radiation
• Information technologies & instrumentation

Practices:

• Asking questions & defining problems
• Developing and using models
• Planning & carrying out investigations
• Analyzing & interpreting data
• Using math & computational thinking
• Constructing explanations & designing solutions
• **Engaging in argument from evidence**
• Obtaining, evaluating, and communicating information

Crosscutting Concepts:

• Patterns
• Cause and effect
• Scale, proportion, and quantity
• Systems & system models
• Energy & matter
• Structure & function
• Stability & change
THE BEAR ASSESSMENT SYSTEM (BAS) BUILDING BLOCKS

Construct Map

Measurement Model

Items Design

Outcome Space


DEVELOPMENT PROCESS

- Literature Review
- Research Group
- Teacher meetings
- Teachers try out items in class
- Student Interviews
- Trial
- General Administration (including analysis)
RESEARCH QUESTIONS

1. What is the nature of the learning progression in the content domain of Structure of Matter?

2. What is the nature of the learning progression in students’ ability to argue from evidence in the domain of Structure of Matter and also, generally?

3. What is the inter-relationship between students’ ability to argue from evidence and their domain specific knowledge? In particular to what extent do the two co-vary?
RESULTS

• Paper-and-pencil format test in the spring of 2013
• 1086 grade 8 students took 1 of 4 forms.
• 4 constructs were explored for this analysis: MAC, PHS, EPC, and ECC
• 128 total items related to one of these 4 constructs.
RESULTS

• Measurement Model: Between-item multidimensional model (Wang, Wilson & Adams, 1997)

\[
\log \left( \frac{P(X_i = j | \theta_d)}{P(X_i = j - 1 | \theta_d)} \right) = \theta_d - \delta_{ij}
\]

• \(X_i\): Response to item \(i\)
• \(\theta_d\): person ability on dimension \(d\)
• \(\delta_{ij}\): the difficulty for step \(j\) of item \(i\)

*Delta Dimensional Alignment (DDA; Schwartz & Ayers, 2011) was applied.*
## RESULTS

**Correlations across dimensions**

<table>
<thead>
<tr>
<th></th>
<th>MAC</th>
<th>PHS</th>
<th>EPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC</td>
<td>1.00</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>PHS</td>
<td>0.93</td>
<td>1.00</td>
<td>--</td>
</tr>
<tr>
<td>EPC</td>
<td>0.82</td>
<td>0.81</td>
<td>1.00</td>
</tr>
<tr>
<td>ECC</td>
<td>0.84</td>
<td>0.84</td>
<td>0.88</td>
</tr>
</tbody>
</table>
RESEARCH QUESTION 1: SUMMARY

Preliminary Results

- In general, there appears to be a progression.
- Additional research is needed to understand why ECC does not appear to be more difficult than EPC.
The history of science is the history of vision and argument

Idea

Phenomena

- Read it
- Write it
- Talk it
- Draw it
- Do it
Knowing why you are wrong matters as much as knowing why you are right!
The Role of Argument in Science
RESEARCH QUESTIONS

1. What is the nature of the learning progression in the content domain of Structure of Matter?

2. What is the nature of the learning progression in students’ ability to argue scientifically in the domain of Structure of Matter and also, generally?

3. What is the inter-relationship between students’ ability to reason scientifically and their domain specific knowledge? In particular to what extent do the two covary?
BEAR Measurement Model

- Measurement Model
- Construct
- Items Design
- Outcome Space
Table 1

Practice progression for evidence-based explanations

1. Student is either provided with evidence and asked to choose the appropriate claim OR student is provided with a claim and asked to choose the appropriate evidence
2. Student is provided with a scientific question, and asked to make a claim and back it up with evidence (the process is scaffolded with hints about the core ideas and prompts for including the claim and evidence)
3. Student is provided with a scientific question, and asked to make a claim and back it up with evidence (the process is scaffolded with hints about the core ideas)
4. Student is provided with a scientific question, and asked to make a claim and back it up with evidence (the process is not scaffolded)
5. Student is provided with a scientific question, and asked to construct a scientific explanation (including a claim, evidence and reasoning) (the process is scaffolded with hints about the core ideas and prompts for including the claim, evidence, and reasoning)
6. Student is provided with a scientific question, and asked to construct a scientific explanation (including a claim, evidence and reasoning) (the process is scaffolded with hints about the core ideas)
7. Student is provided with a scientific question, and asked to construct a scientific explanation (including a claim, evidence and reasoning) (the process is not scaffolded)
<table>
<thead>
<tr>
<th>Description of the Level</th>
<th>Toulmin (1958)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0</td>
<td>Non-scientific</td>
</tr>
<tr>
<td>Level 1</td>
<td>Scientific claim</td>
</tr>
<tr>
<td></td>
<td>Claim</td>
</tr>
<tr>
<td>Level 2</td>
<td>Coordination between claim and evidence</td>
</tr>
<tr>
<td></td>
<td>Claim + data</td>
</tr>
<tr>
<td>Level 3</td>
<td>Reasoned coordination between claim and evidence</td>
</tr>
<tr>
<td></td>
<td>Claim + data + warrant/backing</td>
</tr>
<tr>
<td>Level 4</td>
<td>Modified, reasoned coordination between claim and evidence</td>
</tr>
<tr>
<td></td>
<td>Claim + data + warrant/backing + qualifier</td>
</tr>
<tr>
<td>Level 5</td>
<td>Conditional, modified, reasoned coordination between claim and evidence</td>
</tr>
<tr>
<td></td>
<td>Claim + data + warrant/backing + qualifier + conditions of rebuttal</td>
</tr>
</tbody>
</table>

## A Learning Progression for Argumentation

<table>
<thead>
<tr>
<th>Level</th>
<th>Constructing</th>
<th>Critiquing</th>
<th>Description</th>
<th>Representation of elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td>No evidence of facility with argumentation.</td>
<td></td>
</tr>
<tr>
<td>0a</td>
<td>Stating a claim</td>
<td></td>
<td>Student states a relevant claim.</td>
<td>$C$</td>
</tr>
<tr>
<td>0b</td>
<td></td>
<td>Identifying a claim</td>
<td>Student identifies another person’s claim.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>0c</td>
<td>Providing evidence supporting a claim</td>
<td></td>
<td>Student supports a claim with a piece of evidence.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>1a</td>
<td>Constructing a warrant that links claim and evidence</td>
<td></td>
<td>Student constructs an explicit warrant that links their claim to evidence.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>1b</td>
<td></td>
<td>Identifying a warrant</td>
<td>Student identifies the warrant provided by another person.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>1c</td>
<td>Constructing a complete argument.</td>
<td></td>
<td>Student makes a claim, selects evidence that supports that claim, and constructs a synthesis between the claim and the warrant.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>1d</td>
<td>Providing an alternative counter argument</td>
<td></td>
<td>Student offers a counterargument as a way of rebutting another person’s claim.</td>
<td>$C \rightarrow W \rightarrow E$ $C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>2a</td>
<td></td>
<td>Providing a counter-critique</td>
<td>Student critiques another’s argument. Fully explicates the claim that the argument is flawed and justification for why that argument is flawed.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>2b</td>
<td>Constructing a one-sided comparative argument</td>
<td></td>
<td>Student makes an evaluative judgment about the merits of two competing arguments and makes an explicit argument for the value of one argument. No warrant for why the other argument is weaker.</td>
<td>$C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>2d</td>
<td>Providing a two-sided comparative argument</td>
<td></td>
<td>Student makes an evaluative judgement about two competing arguments and makes an explicit argument (claim + justification) for why one argument is stronger and why the other is weaker (claim + justification).</td>
<td>$C \rightarrow W \rightarrow E$ $C \rightarrow W \rightarrow E$</td>
</tr>
<tr>
<td>2e</td>
<td>Constructing a counter claim with justification</td>
<td></td>
<td>This progress level marks the top anchor of our progress map. Student explicitly compares and contrasts two competing arguments, and also constructs a new argument in which they can explicitly justify why it is superior to each of the previous arguments.</td>
<td>$E \rightarrow C \rightarrow W \rightarrow E$ $C \rightarrow W \rightarrow E$ $C \rightarrow W \rightarrow E$</td>
</tr>
</tbody>
</table>
ITEM DEVELOPMENT PROCESS

PROGRESS MAP TARGET AREA

WRITE ITEM COMMENSURATE WITH PROGRESS MAP

ArgUMENTATION LEARNING PROGRESSION

PEER REVIEW OF FACE VALIDITY

IRT ANALYSIS OF DATA

MODERATION OF ITEM SCORING RUBRICS FOR INTER-RATER RELIABILITY

STUDENT THINK-ALOUDS TO VALIDATE RESPONSE PROCESS VALIDITY

ITEM ADMINISTRATION TO N>600 STUDENTS
Assessment Tasks

www.stanfordargumentation.weebly.com
I agree with Mary because the weight is the same and the sugar would have nowhere to go.

Their teacher asks if they think sugar remains in the water.

Laura says: I think the sugar is gone.

Mary says: I think the sugar is still there.
**LEVEL 0**: Claim and evidence are the fundamental building blocks of argument.

### PROGRESS LEVEL 0

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>INTRINSIC COGNITIVE LOAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stating/identifying an explicit and relevant piece of EVIDENCE</td>
<td><img src="CWE.png" alt="Diagram" /></td>
</tr>
<tr>
<td>Stating/identifying an explicit and relevant CLAIM</td>
<td><img src="C.png" alt="Diagram" /></td>
</tr>
<tr>
<td>No explicit understanding of CLAIMS and/or EVIDENCE</td>
<td><img src="CE.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
LEVEL 1: Claim and evidence must now be coordinated with an explicit warrant.

### PROGRESS LEVEL 1

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>INTRINSIC COGNITIVE LOAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing/Critiquing an explicit and relevant</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td>ARGUMENT or REBUTTAL</td>
<td></td>
</tr>
<tr>
<td>Stating/identifying an explicit and relevant</td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
<tr>
<td>WARRANT</td>
<td></td>
</tr>
</tbody>
</table>
**PROGRESS LEVEL 2**

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>INTRINSIC COGNITIVE LOAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparing relative <strong>SIGNIFICANCE OF</strong> MULTIPLE PIECES OF EVIDENCE</td>
<td>![Diagram]</td>
</tr>
<tr>
<td>Constructing/ Critiquing a <strong>TWO-sided</strong> COMPARATIVE ARGUMENT</td>
<td>![Diagram]</td>
</tr>
<tr>
<td>Constructing/ Critiquing a <strong>ONE-sided</strong> COMPARATIVE ARGUMENT</td>
<td>![Diagram]</td>
</tr>
</tbody>
</table>

**LEVEL 2+: TWO OR MORE EXPLICIT WARRANTS REQUIRED TO COORDINATE INCREASING NUMBERS OF CLAIMS AND EVIDENCE.**
ITEM ANALYSIS

<table>
<thead>
<tr>
<th>logits</th>
<th>Student Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>onions_1</td>
</tr>
<tr>
<td>1</td>
<td>onions_2</td>
</tr>
<tr>
<td>2</td>
<td>onions_3</td>
</tr>
<tr>
<td>3</td>
<td>onions_4</td>
</tr>
<tr>
<td>4</td>
<td>onions_5</td>
</tr>
<tr>
<td>5</td>
<td>onions_6</td>
</tr>
</tbody>
</table>

LEVEL 0

LEVEL 1

LEVEL 2
ASSESSING UNIDIMENSIONALITY

Scree Plot: Eigenvalues from Exploratory Factor Analysis

Yao & Wilson, 2013
REVISION FOR NEXT ROUND OF TESTING:

**Kaela:** I don’t think the onions are why Monica is crying. This is because when you put the onion slices back together, it is the same size as the onion before it was chopped.

---

Monica Townsmad had to chop several onions for dinner. Her two younger sisters, Kaela and Meghan, saw Monica crying as she chopped the first onion. They began to wonder why.

*Kaela:* Why is Monica crying? Is she sad?

*Meghan:* She’s crying because of the onions.

*Kaela:* But the onions are not touching her eyes.

**Kaela:** That can’t be right, because when you put the onion slices back together, it is the same size as the onion before it was chopped. There is no evidence that parts of the onion went into the air.

1. What is Kaela trying to argue?
   Answer Letter___________

2. What is a reason Kaela gives to support her argument?
   Answer Letter___________

3. What is Meghan trying to argue?
   Answer Letter___________

4. What is a reason Meghan gives to support her argument?
   Answer Letter___________

   a) Choppping onions can cause human tears to form.
   b) Monica is sad.
   c) No part of the onion appears to have escaped into the air.
   d) Monica is crying.
   e) Onion parts can get into the air and irritate the human eye.
   f) Monica is not crying because of the onions.
USING EMPIRICAL DATA TO REVISE HYPOTHE Sized MAP
5. Meghan disagrees with Kaela. What might Meghan say to convince Kaela that Kaela’s thinking is wrong?

**Score of 4:** Direct critique to Kaela’s argument. Identifies a flaw in the argument.
- **Example:** “Chemicals released into the air by the air, by the gas, which is not visible, could be the source of the tears.

**Score of 3:** Provided a rebuttal advancing the alternative position.
- **Example:** “The onions juice gets into the air and into the eye causing irritation.”
# Using Empirical Data to Revise Hypothesized Map

## Added to Progress Map: Level 1d

<table>
<thead>
<tr>
<th>Description</th>
<th>Intrinsic Cognitive Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing a COUNTER-ALTERNATIVE argument</td>
<td><img src="Diagram1.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

## Added to Progress Map: Level 2a

<table>
<thead>
<tr>
<th>Description</th>
<th>Intrinsic Cognitive Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing a COUNTER-CRITIQUE</td>
<td><img src="Diagram2.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Average difficulty of items testing scientific argumentation levels

<table>
<thead>
<tr>
<th>Argumentation level</th>
<th>Average item difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0</td>
<td>-1.39</td>
</tr>
<tr>
<td></td>
<td>(Range: -2.59 – 0.31)</td>
</tr>
<tr>
<td>Level 1</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>(Range: -0.62 – 1.04)</td>
</tr>
<tr>
<td>Level 2</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>(Range: 0.91 – 1.56)</td>
</tr>
<tr>
<td>Ability</td>
<td>Item Threshold</td>
</tr>
<tr>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Level 0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>facebook_c_3</td>
<td></td>
</tr>
<tr>
<td>lunch_c_3</td>
<td></td>
</tr>
<tr>
<td>violence_b_2</td>
<td></td>
</tr>
<tr>
<td>facebook_d_2</td>
<td></td>
</tr>
<tr>
<td>lunch_e</td>
<td></td>
</tr>
<tr>
<td>facebook_c_MC</td>
<td></td>
</tr>
<tr>
<td>violence_a</td>
<td></td>
</tr>
</tbody>
</table>
### Automated scoring results for Sugar_e

<table>
<thead>
<tr>
<th>Actual score</th>
<th>Predicted score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>214</td>
</tr>
</tbody>
</table>

Cohen’s kappa = 0.60
What have the data revealed about the difficulty of critique?
Laura says: I think the sugar is gone.

Mary says: I think the sugar is still there.

Who do you agree with more?
Why do you think they are right and why do you think the other student is wrong?
CONSTRUCTING A COUNTER-ALTERNATIVE ARGUMENT

Laura’s claim is incorrect because the sugar’s crystal structure has been dissolve[d] and the sugar molecules are still there.
THINK-ALOUDS

PROVIDING A COUNTER-CRITIQUE

I think I agree with Mary more . . . Laura probably thought that the sugar was gone… She’s just saying that based on what she observed.
THINK-ALOUDS

A picture of the sugar and water particles before they are mixed together is given below.
THINK-ALOUDS

You will be asked to explain why one of the other pictures is not as good as the one you chose. Which one do you want to write about?
When the water and sugar is mixed together and it forms a solution, but it doesn’t happen like that. The sugar and water particles are still individual particles, but they're just mixed together evenly, so you can't depict the sugar particles in the water.

Picture ___ is not as good because...

Coded 49 Excerpts ($\kappa = 0.61$)

CRITIQUE BY ADVANCING ALTERNATIVE POSITION 10%

*It should be spread out more so you can’t see it.*

DIRECT CRITIQUE, NO REASONING 27%

*In C, there’s only one color and it’s white.*

DIRECT CRITIQUE, COMPLETE REASONING 63%

*When the water and sugar is mixed together and it forms a solution, but it doesn’t happen like that. The sugar and water particles are still individual particles, but they're just mixed together evenly, so you can't depict the sugar particles in the water.*
SUMMARY

SUPPORTED PROGRESS MAP FOR ARGUMENTATION

GREATER STUDENT PROPENSITY TO ARGUE IN THE AFFIRMATIVE
– CRITIQUE IS CHALLENGING

MORE STUDENTS WERE ABLE TO CRITIQUE WITH SCAFFOLDING
Issues?

- Structuralist Approach to Argumentation
- True Argumentation is a dialectic
- Inherent problem to the practice model of NGSS
  - What improves?
  - How do we measure it?
RESEARCH QUESTIONS

1. What is the nature of the learning progression in the content domain of Structure of Matter?

2. What is the nature of the learning progression in students’ ability to argue from evidence in the domain of Structure of Matter and also, generally?

3. What is the inter-relationship between students’ ability to argue from evidence and their domain specific knowledge? In particular to what extent do the two co-vary?
PRELIMINARY RESULTS

• Computer format test administered in the spring of 2014
• 278 students (119 8th graders and 159 10th graders)
• Content items here only relate to the EPC construct
• Tasks were developed in alignment with NGSS:
  – Each task includes: Argumentation, content, and “embedded” content
Mark and Kian are discussing what happens when they chop onions. They have two different ideas.

Mark says:
Chopping onions makes me cry because when I cut the onion, some gas is released. The gas goes into the air and gets into my eyes.

Kian says:
I disagree. Chopping onions makes you cry because when the knife slices the onion, some liquid squirts out of the onion and into your eyes.

What is Mark's idea about why people cry when they cut onions?

Mark's idea is that...
Describe the arrangement of molecules in ice, liquid water, and water vapor.

**The arrangement of molecules in ice is...**

- Packed closer together than liquid water and in a repeating pattern
- Spread further apart than liquid water and in a repeating pattern.
- Packed closer together than liquid water and in a random pattern.
- Spread further apart than liquid water and in a random pattern.
Have you ever noticed that when people chop onions they look like they are crying?

In the space below, explain how you think a chemical from the onion could get into a person's eye.

A chemical could get into a person's eye by...
PRELIMINARY RESULTS

• Total of three different “tasks”: Onions, Sugar, and Gases
• Total of 40 items in this analysis
MEASUREMENT MODEL

• Measurement Model: Use both between-item multidimensional models and testlet models (Wang & Wilson, 2005)

\[
\log \left( \frac{P(X_i = j | \theta_d)}{P(X_i = j - 1 | \theta_d)} \right) = \theta_d - \delta_{ij} + \gamma_{d(i)}
\]

• $X_i$: Response to item $i$
• $\theta_d$: person ability on dimension $d$
• $\delta_{ij}$: the difficulty for step $j$ of item $i$
• $\gamma_{d(i)}$: the random effect of testlet $d(i)$ with $\gamma_{d(i)} \sim N(0, \sigma_{\gamma(d(i))}^2)$
UNIDIMENSIONAL MODEL

Science

Item Contexts
S: Sugar
O: Onions
G: Gases

Item Types
A: Argumentation
E: Embedded Content
C: Content

Item SA ... Item SE ... Item SC ... Item OA ... Item OE ... Item OC ... Item GA ... Item GE ... Item GC
TWO-DIMENSIONAL BETWEEN-ITEM MODEL

Item Contexts:
- S: Sugar
- O: Onions
- G: Gases

Item Types:
- A: Argumentation
- E: Embedded Content
- C: Content

Diagram:
- Argumentation
- Content

Subcategories:
- Item SA
- Item SE
- Item SC
- Item OA
- Item OE
- Item OC
- Item GA
- Item GE
- Item GC
THREE-DIMENSIONAL BETWEEN-ITEM MODEL

Argumentation

Embedded Content

Content

Item Contexts
S: Sugar
O: Onions
G: Gases

Item Types
A: Argumentation
E: Embedded Content
C: Content

Item SA ... Item SE ... Item SC ... Item OA ... Item OE ... Item OC ... Item GA ... Item GE ... Item GC
TESTLET MODEL (ONE UNDERLYING DIMENSION)

Science

Item Contexts
S: Sugar
O: Onions
G: Gases

Item Types
A: Argumentation
E: Embedded Content
C: Content

Sugar
Onions
Gases

Item SA... Item SE
Item SC... Item OA
Item OE... Item OC
Item GA... Item GE
Item GC
TESTLET MODEL (TWO UNDERLYING DIMENSIONS)
PRELIMINARY RESULTS

Results for the unidimensional, two-dimensional, three-dimensional, and testlet model

<table>
<thead>
<tr>
<th>Model</th>
<th>Deviance</th>
<th>#Parameters</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unidimensional</td>
<td>13189.68</td>
<td>62</td>
<td>13313.68</td>
<td>13538.59</td>
</tr>
<tr>
<td>2-dimensional</td>
<td>13133.04</td>
<td>64</td>
<td>13261.04</td>
<td>13493.21</td>
</tr>
<tr>
<td>3-dimensional</td>
<td>13049.17</td>
<td>67</td>
<td>13183.17</td>
<td>13426.22</td>
</tr>
<tr>
<td>Testlet – 1 dim</td>
<td>13025.73</td>
<td>65</td>
<td>13155.73</td>
<td>13391.53</td>
</tr>
<tr>
<td>Testlet – 2 dim</td>
<td>12961.55</td>
<td>67</td>
<td>13095.55</td>
<td>13338.60</td>
</tr>
</tbody>
</table>
PRELIMINARY RESULTS

Results for the two-dimensional testlet model

<table>
<thead>
<tr>
<th></th>
<th>Argumentation</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAP/PV Reliability</td>
<td>0.66</td>
<td>0.72</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>
RESEARCH QUESTION 3: SUMMARY

Preliminary Results

- The contexts of the items (e.g. “onions”, “gases”) are important for this data.
- Local dependence an important issue in this set of items.
- Initial response to the newly published science education reform.
RESEARCH SUMMARY

Learning Progressions in Middles School Science Instruction & Assessment (LPS)

• Developed learning progressions for:
  • Structure of Matter
  • Argumentation

• Started investigating the relationship between progressions
Questions and Comments

Contact Information:
Linda Morell: lindamorell@berkeley.edu
Tina Chiu: tchiu129@berkeley.edu
Mark Wilson: markw@berkeley.edu
Jonathan Osborne: osbornej@stanford.edu
Andrew Wild: awild@stanford.edu