Blending Evidence-Centered Design and Universal Design for Learning in Next-Generation Science Assessment

National Council on Measurement in Education

April 2019 | Toronto

The contents of this paper were developed under a grant from the Department of Education. However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

I-SMART

Innovations in Science Map, Assessment & Report Technologies

Participants

Presenters from ATLAS, University of Kansas:

- Meagan Karvonen
- Lori Andersen
- Russell Swinburne Romine
- Gail Tiemann

Discussant: Brian Gong, Center for Assessment

Project Partners

- Maryland – Lead State
- Missouri
- New Jersey
- New York
- Oklahoma
- CAST
- BYC Consulting

Overview of NGSS-Aligned Learning Map Models and Assessment Design Considerations

Meagan Karvonen and Gail Tiemann

ATLAS, University of Kansas

I-SMART

Innovations in Science Map, Assessment & Report Technologies

What prompted this project?

- Persistent challenges with science achievement, pursuit of STEM majors and careers
- Historical approach to teaching science: facts rather than deep connections (NGSS Lead States, 2013)

What prompted this project?

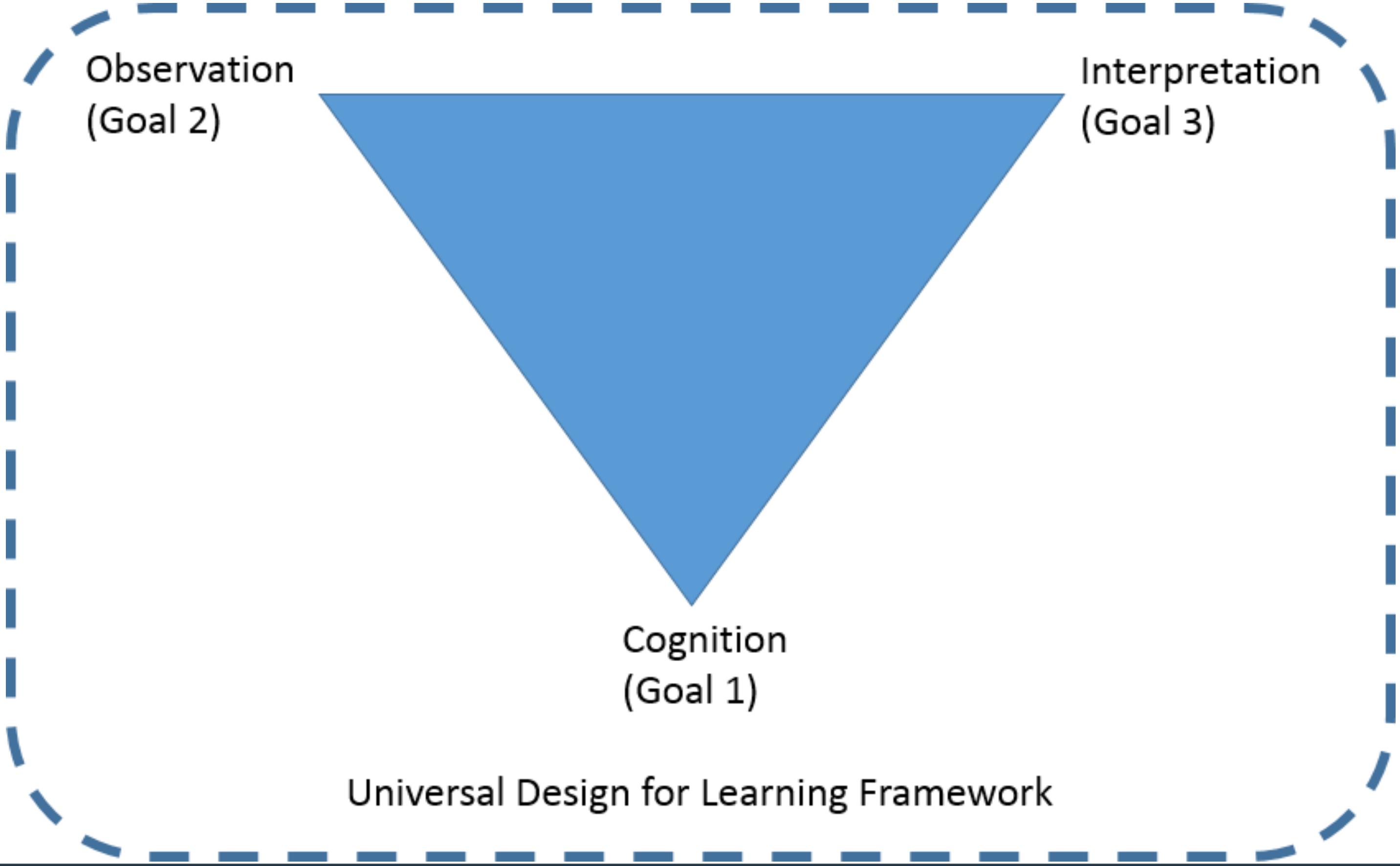
- Shift to the K-12 Framework and NGSS → Multidimensional performance expectations
- Gap between existing science education models and what is needed for new expectations

What prompted this project?

- How to assess students on more complex performance expectations without introducing barriers?
  - Especially a concern for struggling learners and those with disabilities
- How to make assessment results useful and actionable for teachers?

I-SMART Project Purpose

Improve achievement of multidimensional science standards for **students with and without disabilities** through accessible, learning map model-based assessments and reporting tools


Observation
(Goal 2)
Interpretation
(Goal 3)
Cognition
(Goal 1)
Universal Design for Learning Framework

Project Goals

Goal 1 – Develop & evaluate science learning map model

- Expand existing DLM science neighborhoods based on literature reviews
- Connect to math, ELA & pre-academic foundational skills
- Revise map neighborhoods based on external reviewer feedback

Goal 2 – Design, develop & evaluate assessments

- Measure science disciplinary content & engineering practices focusing on the most important KSUs
- Use highly engaging, universally designed, technology-delivered formats
- Revise testlets & concept maps based reviews and pilot test

Project Goals

Goal 3 – Design, develop & evaluate a dashboard

- Develop reporting dashboard, online system, and content that describes student performance on science assessments
- Gather feedback from teacher focus groups and interviews

Goal 4 – Dissemination

- Distribute range of materials to stakeholders including appropriate organizations, educators, and policy makers

Remaining Papers

Share implementation and evaluation (so far) in goals 1-2

- Map design and evaluation
- Assessment design using UDL and ECD principles
- Evaluating new assessment features via cognitive labs

Designing and Evaluating Accessible Science Learning Map Models

Lori Andersen, Meagan Karvonen, & Russell Swinburne Romine

Accessible Teaching, Learning, & Assessment Systems,
University of Kansas

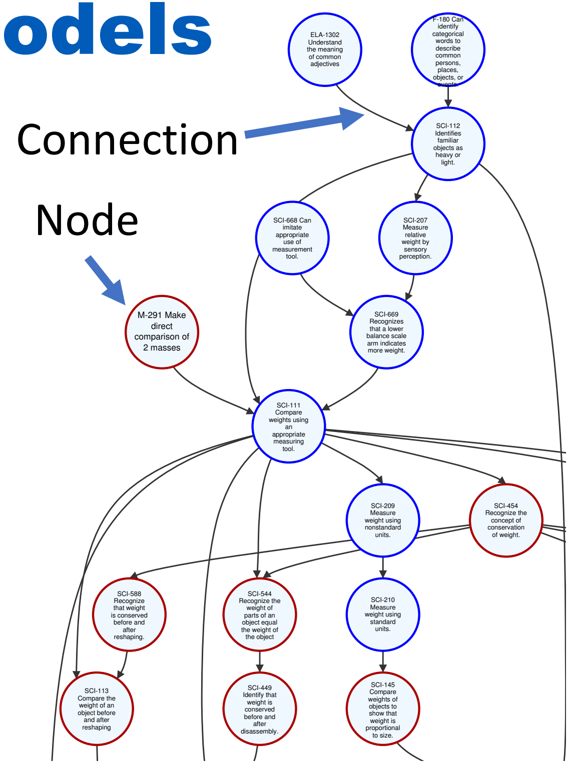
NCME 2019 Conference, Toronto, ON, Canada

Presentation Overview

- Learning Map Models
- Map Design Modifications
- Panel Review Process Refinements
- Post-Panel Review Process
- Target Selection

Background - Learning Map Models

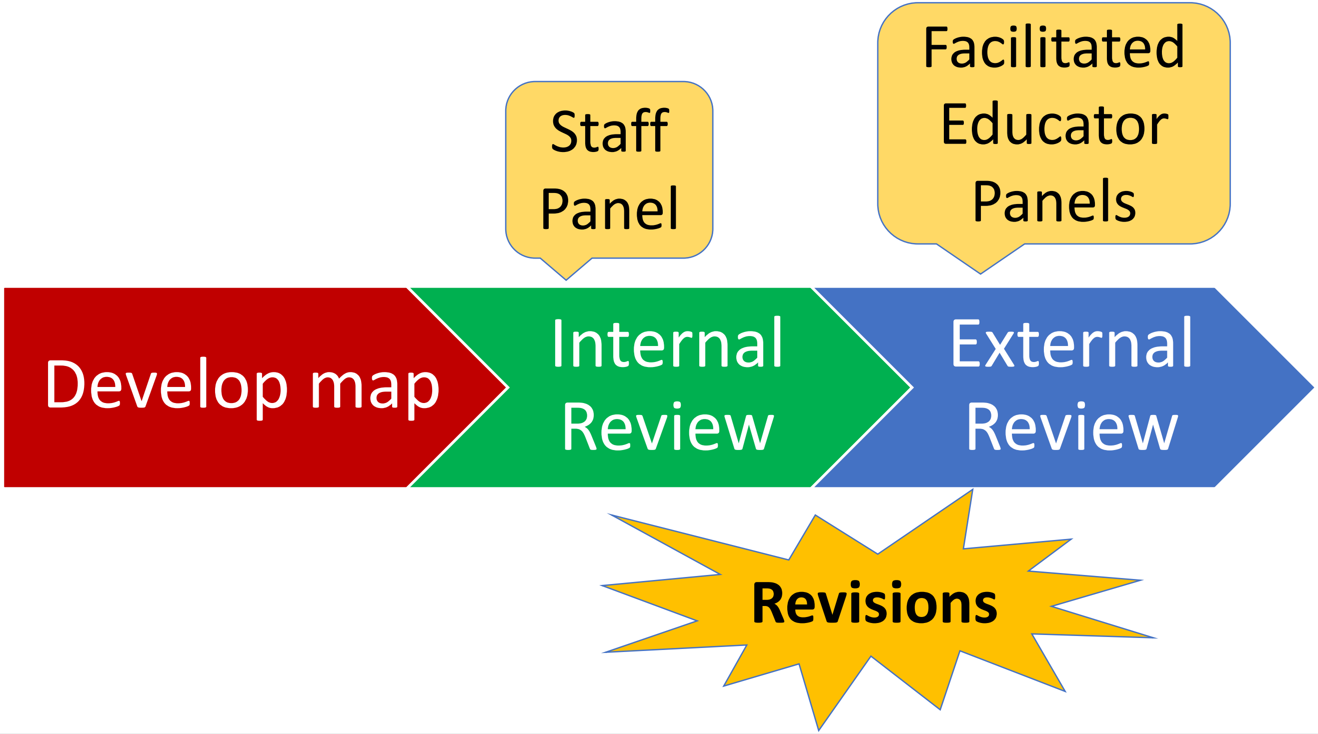
Cognitive models that are fine-grained representations of the knowledge, skills, and understandings required to reach a learning target, which have interconnected, multiple pathways. (Bechard et al., 2012; Kingston et al., 2017)



Map Design Differences from Prior Work

- Maps are **multidimensional,** with additional node types and more multiple pathways
- Maps have **smaller unit size**, with a *neighborhood* for each standard comprised of a disciplinary core idea, science and engineering practice, and crosscutting concept
- Maps connect to ELA, math, and foundation maps

Development Process Overview



Development Process

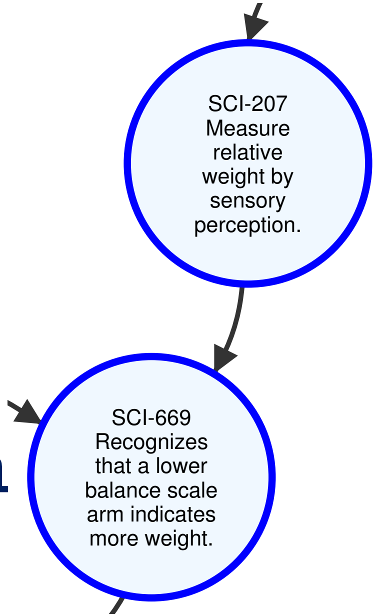
- For each *neighborhood*,
  - Use NGSS resources to identify
    - concepts & skills to include
    - upper boundaries
  - Search and synthesize literature
  - Develop research narrative
  - Create nodes and connections

Design Criteria for Nodes

- Nodes
  - Clear relationship with the standard
  - Unique
  - Small grain size
  - Universal design guidelines

Design Criteria for Connections

- Connections
  - Logical and accurate
  - Appropriate for students with SCD
  - Appropriate for students with sensory, mobility, or communication barriers



Internal and External Reviews

- Internal staff panel evaluated design criteria
    - Individual and consensus determinations
- External facilitated panels of science and special educators
  - 2 evaluations
    - Individual neighborhood
    - Between and among maps

Types of External Review Evaluations

- Individual neighborhoods
  - Node and connection design criteria
  - Global neighborhood criterion
  - Major pathway (prioritize targets)
- Between and among neighborhoods
  - Horizontal criteria
  - Vertical criteria

Panel Review Summary Data – Final Consensus Recommendations on Nodes and Connections by Neighborhood

Neighborhood	Nodes that met criteria	%	Connections that met criteria	%
EE Name	##	%	##	%
Totals	##	%	##	%

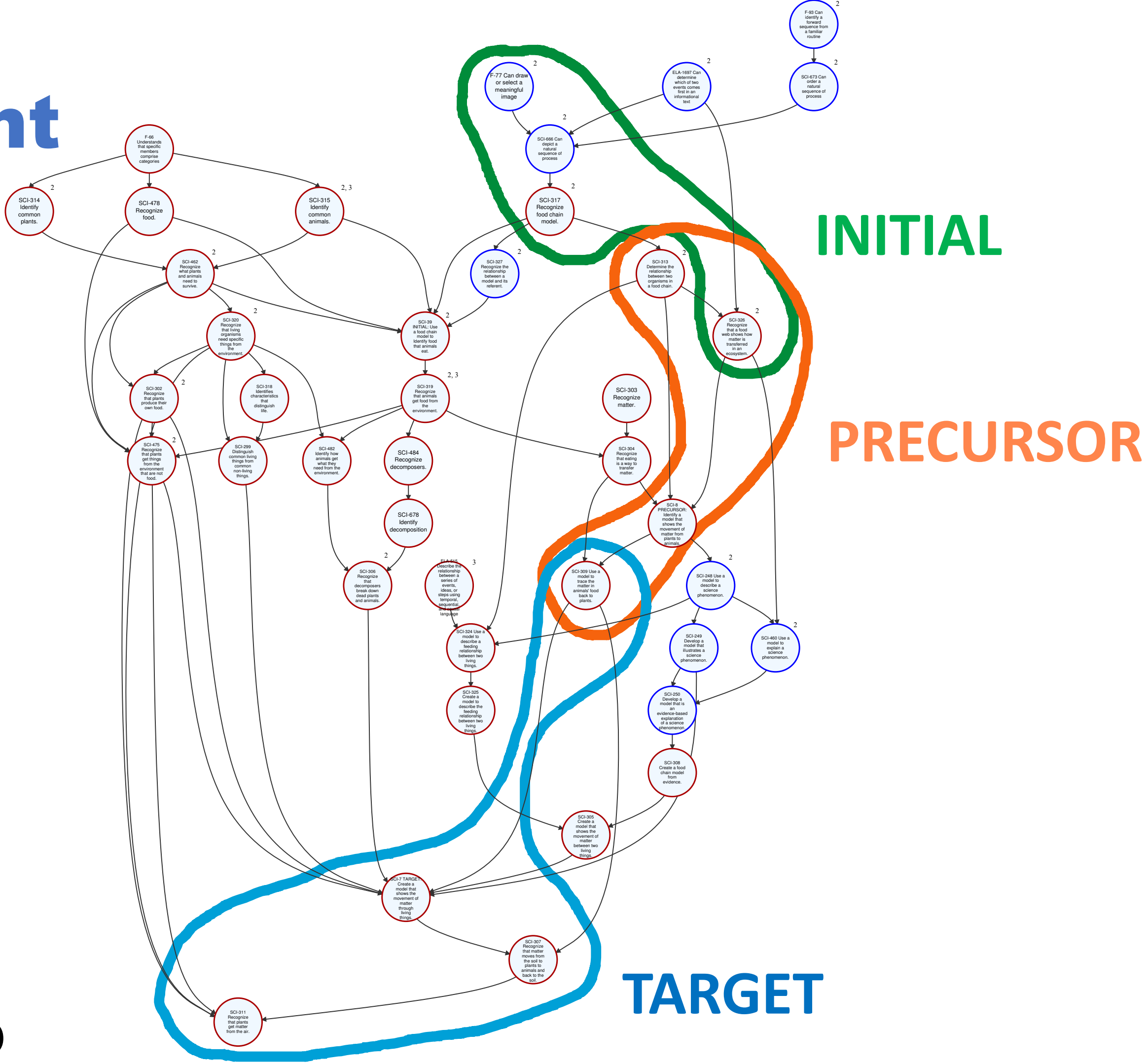
Post-Panel Review Process

- New process
  - Step 1 – staff member accepts panel recommendations that are logical and consistent
  - Step 2 – staff panel discussion
    - Accept or Reject

Identifying Assessment Targets

- Design guidelines for linkage levels
  - Use major pathway
  - Span map for wide range of access points
  - Support future data modeling
- Content Team decisions for including nodes
  - Include DCI and SEP
  - Different complexity

Assessment Target Example



(ATLAS, 2018)

Conclusions

- New map design addresses multidimensionality
- New processes strengthen validity evidence for maps
- Innovative assessment targets increase range of access points
- Empirical evidence is needed

I-SMART

Assessment Design: Integrating Evidence-Centered Design and Universal Design for Learning

Russell Swinburne Romine
Gail Tiemann
ATLAS - University of Kansas
April 6, 2019
10:25 – 11:55 am

I-SMART

Innovations in Science Map, Assessment & Report Technologies

Overall project purpose: Improve achievement of multidimensional science standards for students with and without disabilities through accessible, learning map model-based assessments and reporting tools.

Goal 2 – Design, develop & evaluate assessments

- Measures science disciplinary content & engineering practices focusing on the most important KSUs
- Uses highly engaging, universally designed, technology-delivered formats
- Revises testlets & concept maps based reviews and pilot test

Science Assessments for All Students

- Designing science assessments to engage higher-order thinking without increasing barriers
- Innovative design approaches are needed to develop science assessments linked to the NGSS that are accessible to all students

Overview

- Evidence-Centered Design (ECD: Mislevy, Steinberg & Almond, 2003)
- Universal Design for Learning (UDL: CAST, 2011)
  - Innovative Assessment Design Features
- The Essential Element Concept Map (EECM)

Evidence-Centered Design

- Evidence-centered design framework
  - Provides a systematic, thoughtful mechanism for answering questions regarding student knowledge, evidence, and interpretation of the evidence.

Universal Design for Learning

- Universal Design Guidelines provide students with
    - Multiple means of engagement
    - Multiple means of representation
    - Multiple means of action and expression

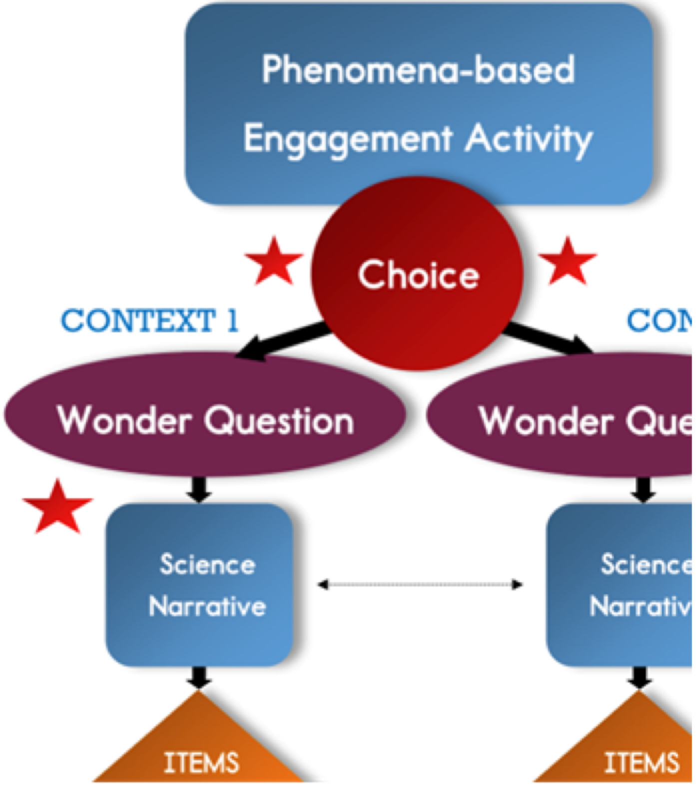
The Universal Design for Learning Guidelines

CAST | Until learning has no limits

	Provide multiple means of **Engagement**  Affective Networks The "WHY" of Learning	Provide multiple means of **Representation**  Recognition Networks The "WHAT" of Learning	Provide multiple means of **Action & Expression**  Strategic Networks The "HOW" of Learning
Access	Provide options for **Recruiting Interest** (7) • Optimize individual choice and autonomy (7.1) • Optimize relevance, value, and authenticity (7.2) • Minimize threats and distractions (7.3)	Provide options for **Perception** (1) • Offer ways of customizing the display of information (1.1) • Offer alternatives for auditory information (1.2) • Offer alternatives for visual information (1.3)	Provide options for **Physical Action** (4) • Vary the methods for response and navigation (4.1) • Optimize access to tools and assistive technologies (4.2)
Build	Provide options for **Sustaining Effort & Persistence** (8) • Heighten salience of goals and objectives (8.1) • Vary demands and resources to optimize challenge (8.2) • Foster collaboration and community (8.3) • Increase mastery-oriented feedback (8.4)	Provide options for **Language & Symbols** (2) • Clarify vocabulary and symbols (2.1) • Clarify syntax and structure (2.2) • Support decoding of text, mathematical notation, and symbols (2.3) • Promote understanding across languages (2.4) • Illustrate through multiple media (2.5)	Provide options for **Expression & Communication** (5) • Use multiple media for communication (5.1) • Use multiple tools for construction and composition (5.2) • Build fluencies with graduated levels of support for practice and performance (5.3)
Internalize	Provide options for **Self Regulation** (9) • Promote expectations and beliefs that optimize motivation (9.1) • Facilitate personal coping skills and strategies (9.2) • Develop self-assessment and reflection (9.3)	Provide options for **Comprehension** (3) • Activate or supply background knowledge (3.1) • Highlight patterns, critical features, big ideas, and relationships (3.2) • Guide information processing and visualization (3.3) • Maximize transfer and generalization (3.4)	Provide options for **Executive Functions** (6) • Guide appropriate goal-setting (6.1) • Support planning and strategy development (6.2) • Facilitate managing information and resources (6.3) • Enhance capacity for monitoring progress (6.4)

UDL Features Embedded into I-SMART Testlets

- Phenomena-based engagement
- Student Choice
- Wonder Questions
- Science Narratives
- Embedded Items



UDL in an Evidence-Centered Design Framework

- **Essential Element Concept Map** is a document that specifies the connection between the **content**, a testlet's **design elements**, and student **observations**.

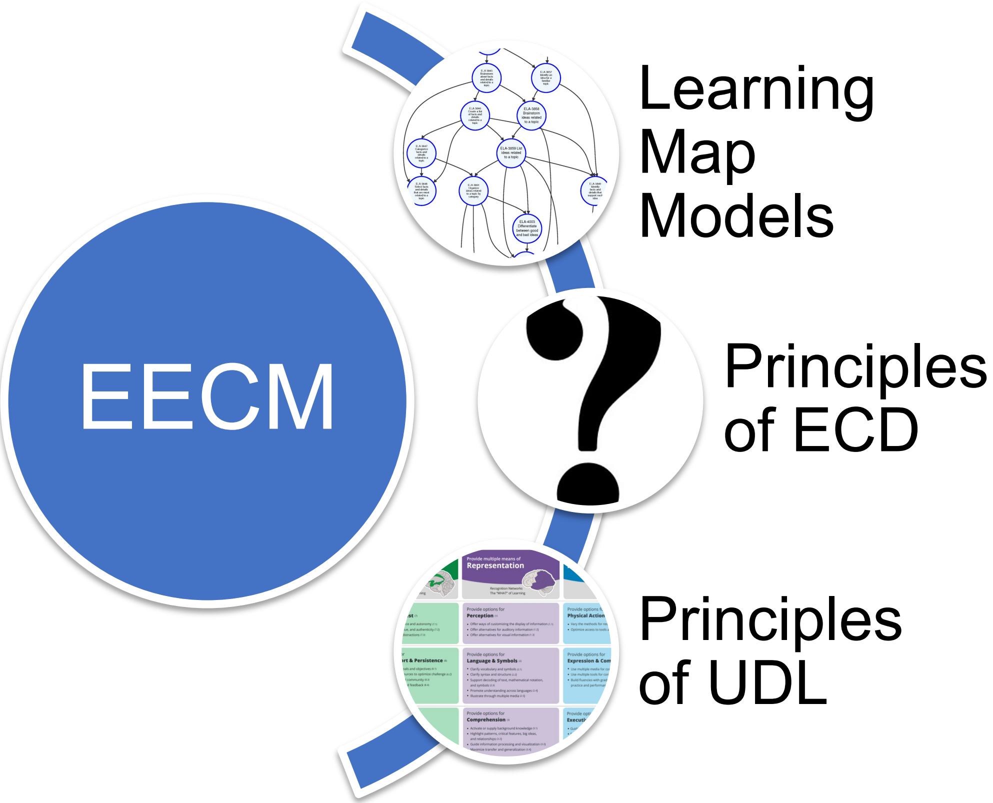
- Provides guidance to item-writers in developing testlets and incorporating UDL options.

Essential Element Concept Maps

- **Essential Element Concept Map** is a document that specifies the connection between the **content**, a testlet's **design elements**, and student **observations**. (DLM, 2016, Bechard, et al., in press)
- Supports the development of well-aligned items
- Leverages the value of a theory-grounded, intentional design process in an easy to use visual format

The EECM

- Built around a content standard
- Designed as a tool for item writers to integrate multiple frameworks



Building on the work from DLM

- EECMs Include:
  - Prerequisite and requisite skills
  - Common misconceptions
  - Key vocabulary
  - Common questions to ask students
  - Level descriptions that focus on how students can demonstrate understanding
  - Specific statements of how DCI and SEP are conceptualized in each assessment target

Example EECM Sections

EE.HS.LS2-2

Essential Questions for the Initial level

- Does the student understand that different objects can be members of the same category?
- Can the student identify common animals?
- Does the student recognize that different members of one type of organism comprise a population?
- Does the student recognize that food and shelter are needed for survival?

Initial Level Name	Initial Level Description
EE.HS.LS2-2.I	Identify common animals and populations, recognize their survival needs.

Vocabulary		Misconceptions
Concepts	habitat needs of a species	(F-66) The student does not distinguish categories of living things, such as animals. (SCI-315) The student cannot identify common animals. (SCI-527) The student does not recognize population. The student includes more than one type of organism when determining a population rather than only counting members of one species. (SCI-501) The student does not recognize that food and shelter are needed for survival. The student indicates nonessential items as needed for survival. The student confuses the scientific usage of the word food with its common usage.
Words	food, shelter, survive	

Information	Phenomenon	Wonder Question	Science and Engineering Practices
Plants and animals are categories of living things (LS2.A) → Identify common animals (LS2.A) → Recognize that groups of the same kind of living things live in the same area (population; LS2.A) → Recognize food and shelter as needed for survival (LS2.A)	General Mechanism – Organisms live together as populations in ecosystems. Example: A group of rabbits lives in the forest. Trees are in the forest.	Not used at the initial level.	Mathematics and computational thinking Students can count animals in a population. They understand how to use numbers with meaning and can use simple graphs to compare quantities or notice patterns.

Example of EECM Sections

Target linkage level nodes

Nodes should be presented in the testlet in an order that creates a logical flow in inquiry activity, which may differ from the order that they occur in the map.

Nodes (order from map)	**Description**	**Observation & Example Questions to Ask**
SCI-309 Use a model to trace matter in animals' food to plants. ↓	Linking node Integrated Node 4 items Use a model to trace the matter in animals' food back to plants.	The student is presented with a simple food web (e.g., grass -> rabbit -> fox). The student identifies that the matter in the fox's food came from grass. Example Questions: What does the model show about how the [organism] gets matter? Which model shows how [organism] gets matter?
SCI-7 Create a model that shows the movement of matter through living things ↓	Integrated Node 3 items Create a model that shows the movement of matter (e.g. plant growth, eating, composting) through (three or more) living things.	The student is shown a partially complete food chain model (e.g., one organism or arrow is missing). The student is asked to fill in the missing item based on the description of the feeding relationships from an engagement activity story. Example Questions: Which food chain shows how matter moves? Put the plants and animals in the correct box to show how matter moves [drag and drop item]. What goes between [organism1] and [organism2] to show how matter moves [AOs are types of arrows]?
SCI-307 Recognize that matter moves from the soil to plants to animals and back to the soil. ↓	DCI Node 3 items Recognize that matter moves from the soil to plants to animals and back to the soil.	When shown an example of a cycle food web (e.g., grass -> rabbit -> fox -> worm) the student identifies that food web shows that matter moves from grass to rabbit to fox to worm to soil. The student identifies that plants get nutrients from the soil, but not matter. [Note: Confusing food and nutrients is a misconception. Nutrients for plants are like vitamins for people.] Example Questions: What does the model show about how the [organism] gets matter? What does [character's] food chain show about matter?
SCI-311 Recognize that plants get matter from the air.	DCI Node 3 items Recognize that plants get matter from the air (i.e., carbon dioxide).	When asked, "How does a plant get material it needs to grow?", the student indicates that plants get matter (carbon dioxide) from the air. For example, when asked, "How does a tree get material it needs to grow?", the student indicates that trees take in air through their leaves to get the material they need to grow. Example Questions: What helps a [plant] get matter? How does [plant] get matter to grow? How does a [plant type] take in the material it needs to grow? What is the material that [plant type] uses to grow?

Use of EECMs in Item Writing

Item writers were able to:

- Synthesize information for each linkage level in their assigned Essential Element.
- Narrow their focus and become familiar with the skills and content required by the nodes in their assigned linkage level.
- Choose a phenomenon to explore in their testlet.
- Create a Choice or a Wonder Question.

Additional UDL Guidance for Item Writers

Provide multiple means of engagement

- **Provide options for Recruiting interest**:
  - Phenomenon is a common, high interest situation that a student might experience and makes connections to the real world. Student choice.
- **Provide options for Self-regulation:**
  - Items asking students to reflect on performance develop self-assessment and reflection.

UDL Options for Item Writers

Provide multiple means of representation

- **Provide options for Language & Symbols:**
  - Use analogies to support understanding of concepts. Use video or images to support decoding.
- **Provide options for Comprehension:**
  - Science narrative provides background knowledge, big ideas, and relationships. Represent relationships with diagrams representing only the most relevant information.

UDL Options for Item Writers

- Provide multiple means of action and expression
  - **Provide options for Expression & Communication:**
    - A variety of item response types (e.g., multiple choice, drag and drop) provide multiple tools for construction and composition.
  - **Provide options for Executive function:**
    - Story character thinkalouds in testlets support students' planning and strategy development in science problem solving. Items asking students to reflect on performance enhance students' capacity for monitoring progress.

Examples of UDL Considerations

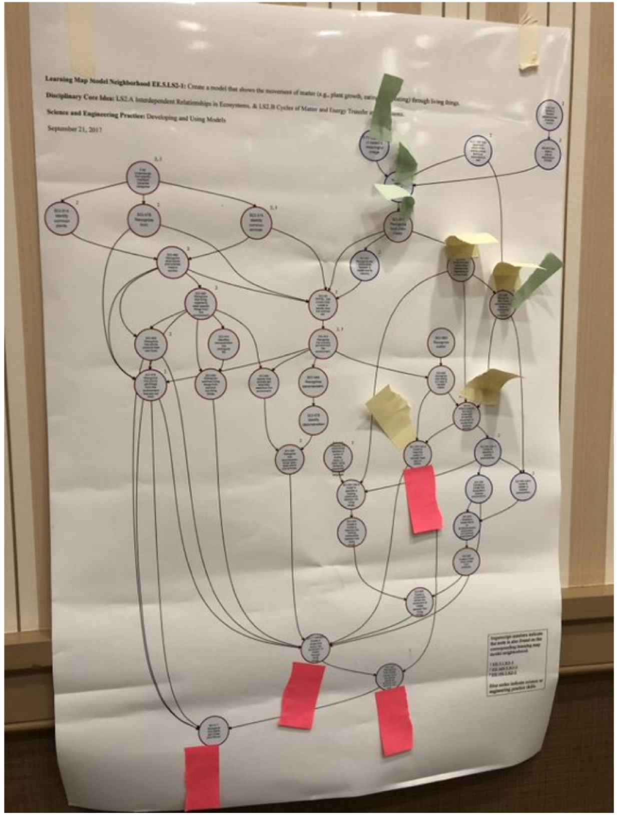
- Example: **Principle for Action & Expression: Executive Function**
  - "What should [character] do next?"
- Example: **Principle for Representation: Language & Symbols**
  - Include a short video of a phenomenon
- Example: **Principle for Engagement: Self-Regulation**
  - Item asking students to reflect on performance: “How did you do?”

The Item Writing Process

- Advance and in-person training
- Using the EECM as a guide
- Peer brainstorming and collaboration
- Storyboarding a testlet
- Peer review
- Drafting content

Training and Resources Supporting UDL

- Learning Map Neighborhood Activity
- EECMs
- Storyboard Organizers
- Item Writing Manual
- UDL Options for Item Writers
- Testlet Template PPTs



Item Writer Survey

- 83% of item writers rated the EECM as a "very effective" tool
- 83% of item writers rated brainstorming with colleagues "very effective"
- 100% of item writers rated feedback from staff as "very effective."

Lessons Learned

- Using the UDL guidelines and checkpoints across the test development process was a valuable tool for staff to self-assess
- Integrating UDL and maps into an ECD based approach requires significant front-end effort
- Item writers respond positively to the EECMs

Using Cognitive Labs to Evaluate Innovative Features of Next Generation Science-Aligned Assessments

Gail Tiemann, PhD

University of Kansas

I-SMART

Innovations in Science Map, Assessment & Report Technologies

Research Questions

1. How do students interact with the features of innovative item types and with innovative testlets?
2. How much time is required to complete a testlet?
3. Do students' responses represent the science performance expectations the items were designed to measure?
4. What are students' and teachers' perceptions of students' experiences with the new testlets?

Prototype Testlets

- Rich science narrative following an inquiry process and a science phenomenon
- Elementary, middle school, and high school grade bands

Prototype Testlets

- One essential element per testlet
- Four learning map model nodes
- 3-4 scored items per node

Features based on UDL Framework

- Choice of Topic
  - Initial and Precursor Linkage Levels Only
  - Construct-relevant or character-based choice
- Media, unscored items to engage interest, self-assessment

Students

- Group 1 – students eligible for Dynamic Learning Maps alternate assessment
- Groups 2 and 3 – students with and without disabilities performing significantly below grade level

Students

- Received instruction on the content
  - Teacher survey probed this information
- Any grade in the grade band
- Initial level – communicate an answer through any response mode

Students

- Precursor and Target Levels
  - Computer-based
- Initial Level
  - Facilitator administers 1:1
  - Answers entered into computer by facilitator
  - Teacher present to assist with administration and interpretation

Data Sources

- Think aloud and retrospective comments, where possible
- Video and audio tape
- Screen recordings for computer-based
- Teacher and student interviews
- Survey probing instruction and mastery of content

Labs Completed

	Initial – Group 1	Precursor – Group 1	Target – Group 1	Target – Group 2/3
Elementary	6	NA	0	0
Middle School	NA	2	2	4
High School	5	NA	0	0

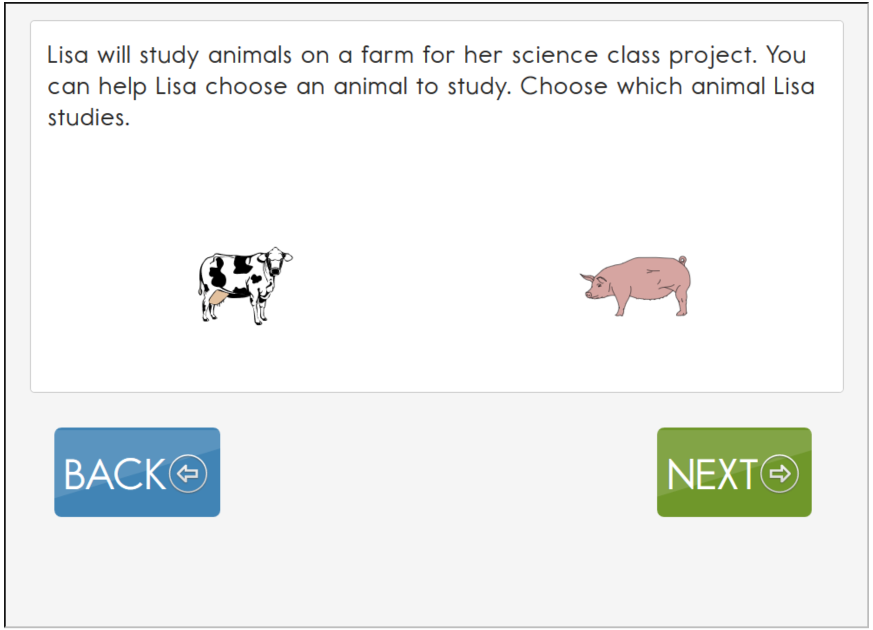
Total – 19 students

States – 2, Schools - 3

Results - RQ 1 Testlet features

Choice

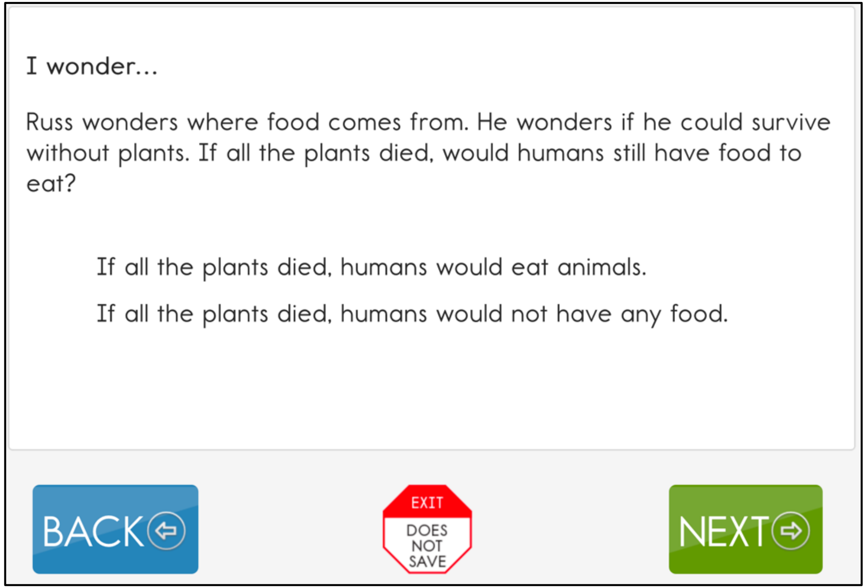
- Initial Level n=11
  - 8 of 11 students made intentional choice
- Precursor Level n=2
  - No difference in student engagement between two different choice options



Results - RQ 1 Testlet features

I Wonder (n=8)

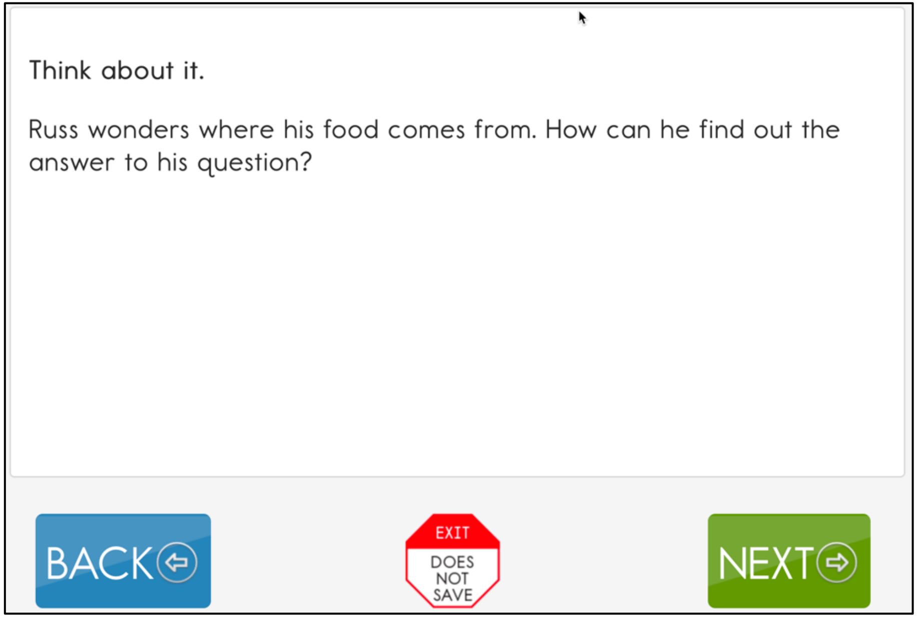
- Middle School Target and Precursor
- Presented twice
- 1 changed correct to misconception
- 3 retained misconception
- 2 changed misconception to correct
- 2 correct both times



Results - RQ 1 - Testlet features

Think About It (n=6)

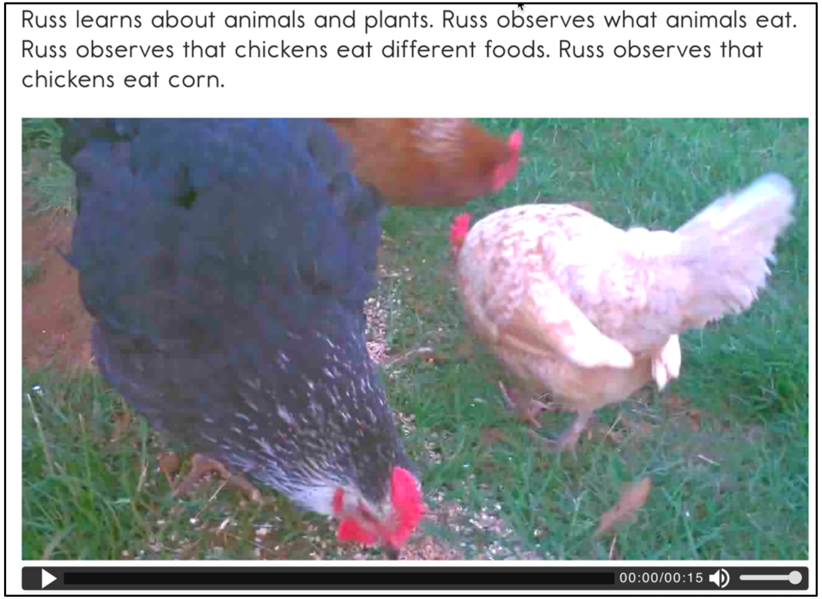
- Middle School Target and Precursor
- Question followed by answer
- Occurs twice in testlet
- Mixed responses
- Second instance, 5 paused to answer out loud



Results - RQ 1 - Testlet features

Video (n=8)

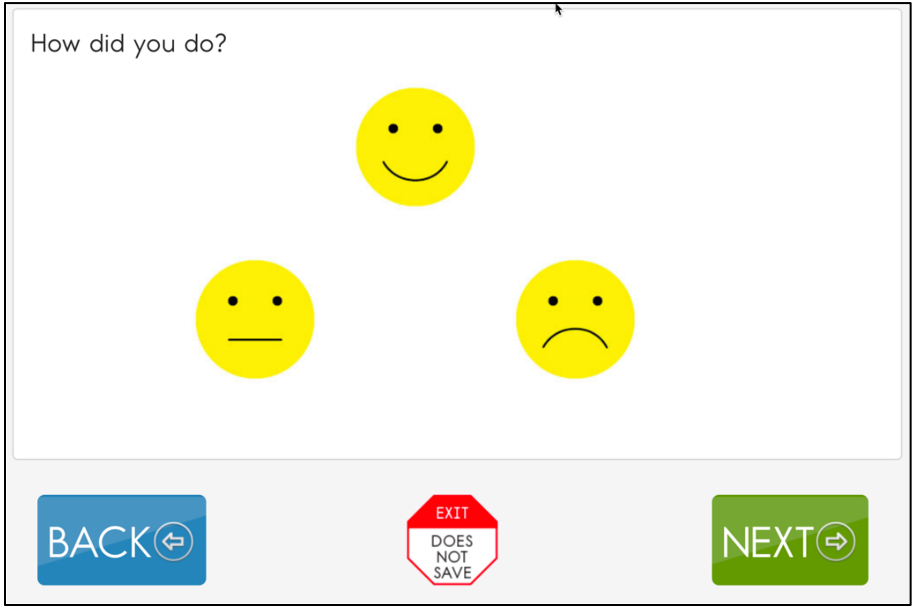
- Middle School Target and Precursor
- Encourage interest, not required for answers
- 6 needed help to play
- Delayed loading startled students
- Tech concerns addressed



Results - RQ 1 - Testlet features

Self-assessment (n=8)

- Middle School Target and Precursor
- All 8 answered ☺



Results – RQ 2 Testlet Time

	Group	N	Item Count	Time Range
Choice-based	1	13	14-17	11:47 – 25:00
Extended Narrative - Target	1	2	16	17:41 – 18:20
Extended Narrative - Target	2/3	4	16	12:21 – 29:28

Middle School students delivered substantial think aloud and retrospective comments.

Results – RQ 3 Content & Performance Expectations

	Group	N	Construct-Relevant Responses	Number of Scored Items
Choice-based-Precursor	1	1	8	14
Extended Narrative - Target	1	2	10,11	14
Extended Narrative - Target	2/3	4	5, 10, 10, 14	14

Analysis based on item specifications – intended response process, misconceptions, guessing, unknown process

Results – RQ 3 Content & Performance Expectations

	Group	N	Students with Evidence of Construct-Relevant Responses
Elementary - Initial	1	6	1
High School - Initial	1	5	3

Construct-irrelevant evidence included picking items based on position, not looking at all choices, random choices. Construct-relevant evidence included teacher interpretations, instruction received, clear answer choices, answer options in variety of positions.

Results – RQ 4 Perceptions

- Length
  - 3 of 8 students too long, 5 just right or normal
  - Teachers did not comment
- Difficulty
  - 3 of 6 students at target-level felt too easy
    - 2 described repetition as a dislike
  - 3 teachers felt content too advanced at initial level
  - Concern about accessibility for students who do not eat

Results – RQ 4 Perceptions

- Media
  - Students liked - 1 suggested more pictures
  - Teacher suggested more realistic, larger pictures
  - Teachers of students at initial level, pictures were unfamiliar
- General Usability
  - Teacher – good flow of content from screen to screen.
  - Some unfamiliar layouts

Discussion

- Results and exploratory and formative
- Six additional interviews completed in March, not included

Discussion

- UDL features were novel, evidence generally suggests features are engaging without adding barriers
  - Difficulty with I Wonder – potential lack of exposure to inquiry
  - Think About It – need more evidence, better probes

Discussion

- Longer than usual tests, but times within acceptable limits
- Students generally interpreting content as intended
  - Students at initial level did make correct selections, especially with more familiar content.
  - Two teachers concerned with difficulty, more item difficultly will be explored during pilot

Discussion

- Students generally liked content
- Media was a favorite, suggestions for improvement addressed in testing platform
- Teacher involvement critical for cognitive lab success, especially at initial level

Discussion

Brian Gong, Center for Assessment

Contact Us…

ismart@ku.edu

ismart.works

785-864-7093

The contents of this presentation were developed under a grant from the Department of Education. However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.