Designing Tasks for Assessing Three-Dimensional Science Learning

Joseph Krajcik (Michigan State University), Christopher Harris (SRI International), Lou DiBello and Jim Pellegrino (University of Illinois, Chicago), Edys Quellmalz and Matt Silberglitt (WestEd)

DRK-12 PI Meeting Panel Session • June 3, 2016 • Washington, DC

The projects showcased in this session are funded by the National Science Foundation, grant numbers 1316903, 1316908, 1316874, and XX) . Any opinions, findings, and conclusions or recommendations at this session or in these materials are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Panelists and Projects

Next Generation
Science Assessment

SimScientists
WestEd

Next Generation Science Assessment

The challenge:

How can we create assessments that integrate the three dimensions of the NGSS and help teachers assess student's progress toward achieving performance expectations?

SRI Education

CREATE for STEM Institute

NSF

UIC LEARNING SCIENCES RESEARCH INSTITUTE

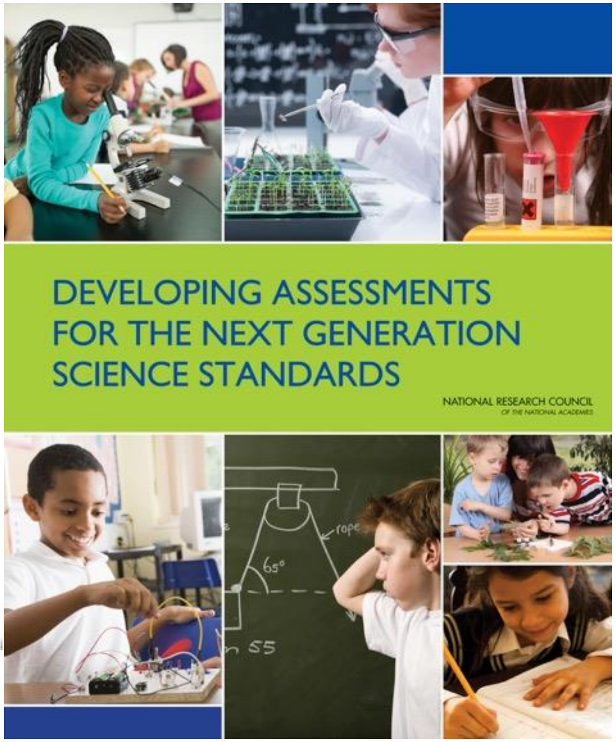
The Concord Consortium

NGSA Project Aims

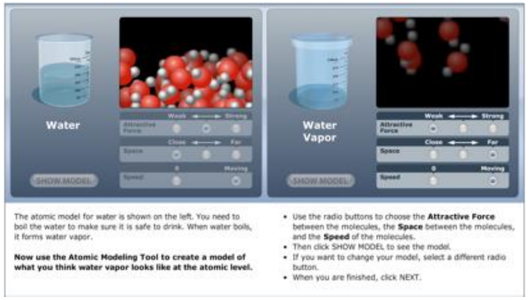
A. Articulate a principled design approach for constructing classroom-based assessments that align to NGSS

B. Use the approach to develop and test technology-based assessment tasks and rubrics (middle school physical science)

C. Engage in a co-design process with science teachers to develop guidelines and strategies for classroom use

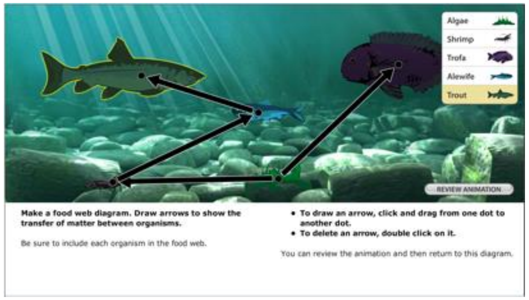


Matter

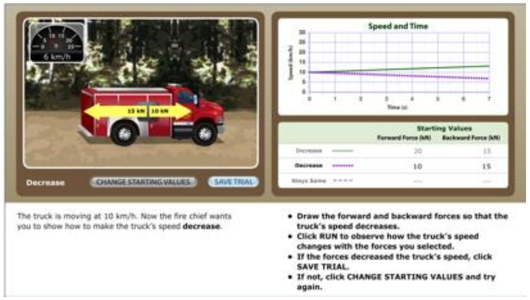


CLASS REPORT

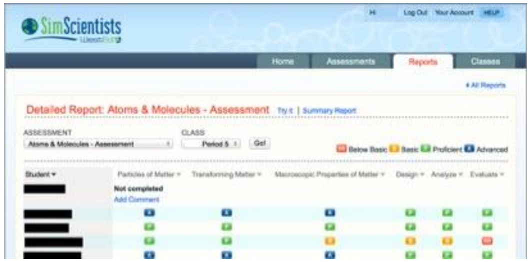
Ecosystems



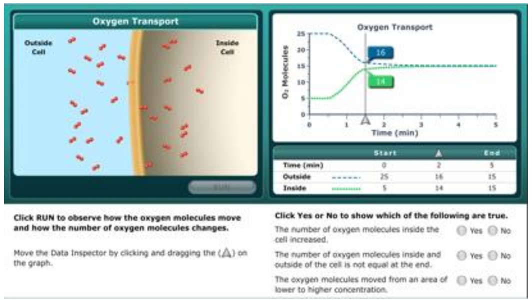
Motion



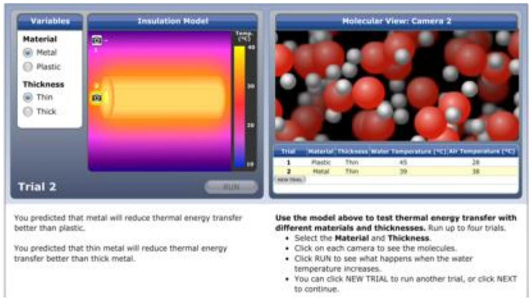
STUDENT DETAIL REPORT



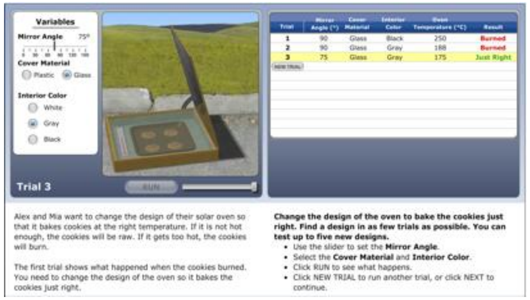
Cells



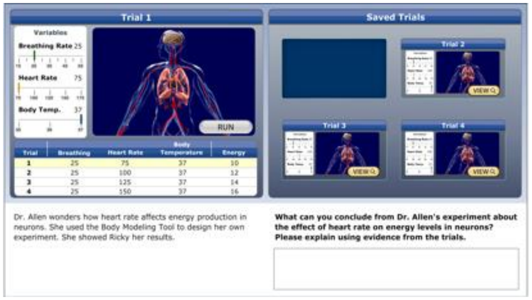
Energy



Waves

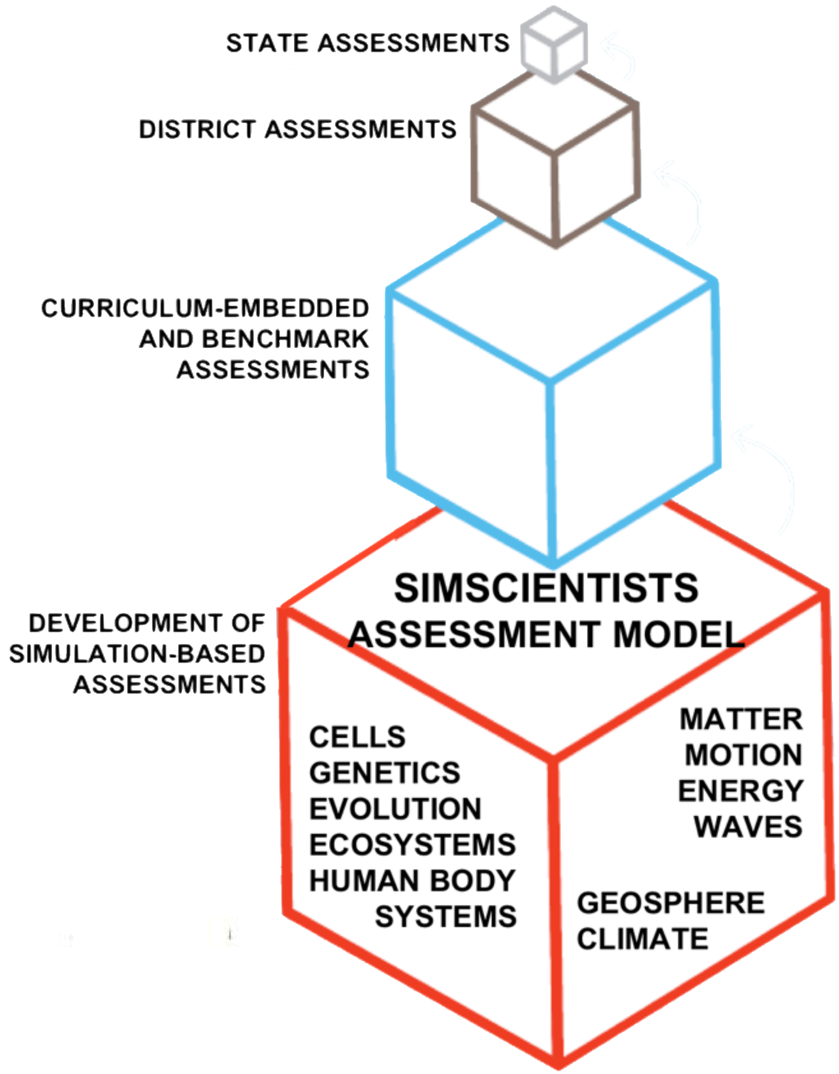


Human Body Systems



SimScientists Project Aims

- Design simulation environments that model dynamic science system phenomena "in action"
- Create simulation-based assessments and curriculum supplements to assess and promote NGSS
- Provide evidence of
  - impacts on science learning
  - technical quality
  - feasibility of implementation
  - potential in balanced state science assessment system



Session Overview

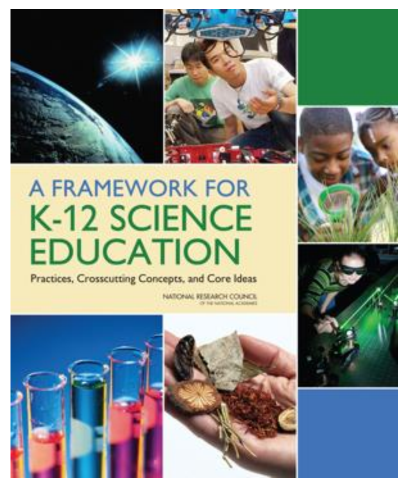
- Vision underlying NGSS
- Challenges in designing NGSS-aligned assessment tasks
- Explore what “three-dimensional” tasks might look like
- Overview of 2 design approaches
- Discussion of the approaches

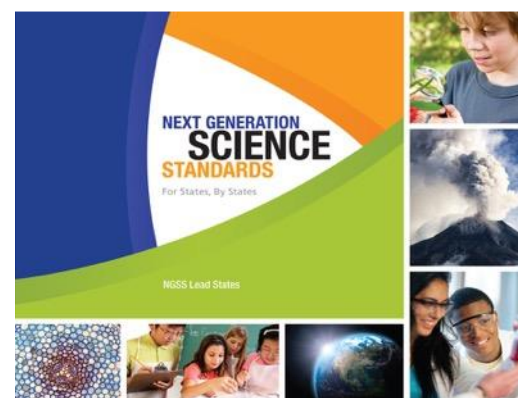
The Vision Behind NGSS

Knowing how to use and apply what you know…

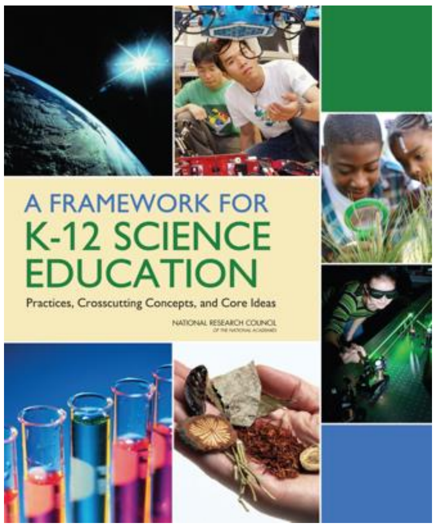
empowers you – in your own learning about the world and your participation in it.

Goal is for every student, from the earliest grades onward, to have coherent and sequenced instruction that provides opportunities to do the “walk and talk” of science and engineering.





What is Really New in the NGSS?

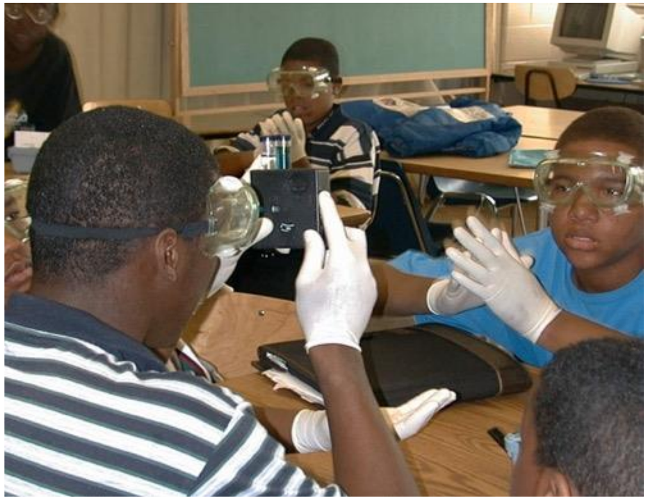


NEXT GENERATION SCIENCE STANDARDS

1. Focus on explaining phenomena or designing solutions to problems
2. Three–Dimensional learning
   - Organized around disciplinary core ideas
   - Use of crosscutting concepts
   - Central role of scientific and engineering practices
3. Standards expressed as performance statements that integrate the 3 dimensions
4. Coherence: building and applying ideas over time and across disciplines
5. Focus on all learners

What is Three-Dimensional Learning?

Three-dimensional learning shifts the focus of the science classroom…



…to where students use disciplinary core ideas, crosscutting concepts with scientific practices to **explore, examine,** and **explain** how and why phenomena occur and to **design solutions** to problems.

How the New Standards are Different

Standards expressed as <u>performance expectations</u>:

- Combine practices, core ideas, and crosscutting concepts into a single statement of ***what is to be assessed***
- Requires students to demonstrate ***knowledge-in-use***
- Performance Expectations are not instructional strategies or objectives for a lesson – ***they describe achievement, not instruction***
- Intended to describe the end-goals of instruction – ***the student performance at the conclusion of instruction***

An NGSS performance expectation

MS-PS1 Matter and its Interactions

Students who demonstrate understanding can:

MS-PS1-2. **Analyze and interpret data on the properties of substances before and after the substances interact to determine if a chemical reaction has occurred.** [Clarification Statement: Examples of reactions could include burning sugar or steel wool, fat reacting with sodium hydroxide, and mixing zinc with hydrogen chloride.] [*Assessment boundary: Assessment is limited to analysis of the following properties: density, melting point, boiling point, solubility, flammability, and odor.*]

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Analyzing and Interpreting Data Analyzing data in 6–8 builds on K–5 and progresses to extending quantitative analysis to investigations, distinguishing between correlation and causation, and basic statistical techniques of data and error analysis. • Analyze and interpret data to determine similarities and differences in findings. - - - - - - - - - - ***Connections to Nature of Science*** **Scientific Knowledge is Based on Empirical Evidence** • Science knowledge is based upon logical and conceptual connections between evidence and explanations.	**PS1.A: Structure and Properties of Matter** • Each pure substance has characteristic physical and chemical properties (for any bulk quantity under given conditions) that can be used to identify it. **PS1.B: Chemical Reactions** • Substances react chemically in characteristic ways. In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants.	**Patterns** • Macroscopic patterns are related to the nature of microscopic and atomic-level structure.

An NGSS performance expectation

MS-PS1 Matter and its Interactions

Science Practice
Analyzing and Interpreting data

Disciplinary Core Idea

PS1.A: Each pure substance has characteristic physical and chemical properties (for any bulk quantity under given conditions) that can be used to identify it.

PS1.B: Substances react chemically in characteristic ways. In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants.

Crosscutting Concept
Patterns

Performance Expectation

MS-PS1-2. Analyze and interpret data on the properties of substances before and after the substances interact to determine if a chemical reaction has occurred.

Assessment Challenges

- *How can we design integrated assessment tasks* in which students make sense of phenomena or design solutions to problems so that they provide evidence of 3-dimensional learning?

- *How do we use performance expectations in order to construct assessment tasks* that can be used during instruction?

- *How do we make these tasks (in)formative* so that they can be used during instruction to help teachers gauge students' progress toward achieving the performance expectations?

NGSA Task Activity

① **Review the two performance expectations and the accompanying assessment task**

① **Discuss with colleagues:**

To what extent does this task provide information on students' building toward the selected PEs?

SimScientists Task Walkthrough

Showcase a Task and describe its intended use

Description of NGSA Design approach

Next Generation
Science Assessment

Assessment design goals

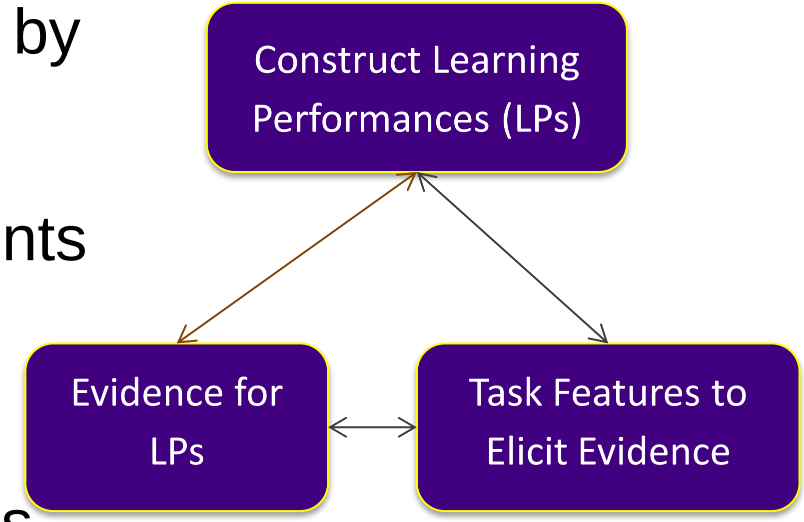
- Tasks aligned with specific **3-dimensional NGSS performance expectations** (middle school physical science and life science)

- Designed for **classroom-based, formative use** to help teachers guide their students toward achieving standards, and to help teachers identify formative assessment opportunities

- Collect and analyze mixed sources of data to **determine validity** of single tasks and groups of tasks, including expert reviews, cognitive laboratory think-aloud studies with students, teacher interviews, classroom observations, and performance studies of groups of tasks with samples of students.

Assessment as an Argument from Evidence: Three Questions

- What do we want students to be able to know and do? (Described by our **Learning Performances**)
- What kinds of evidence will students need to provide to demonstrate proficiency?
- What kinds of tasks / task features will elicit the desired evidence?

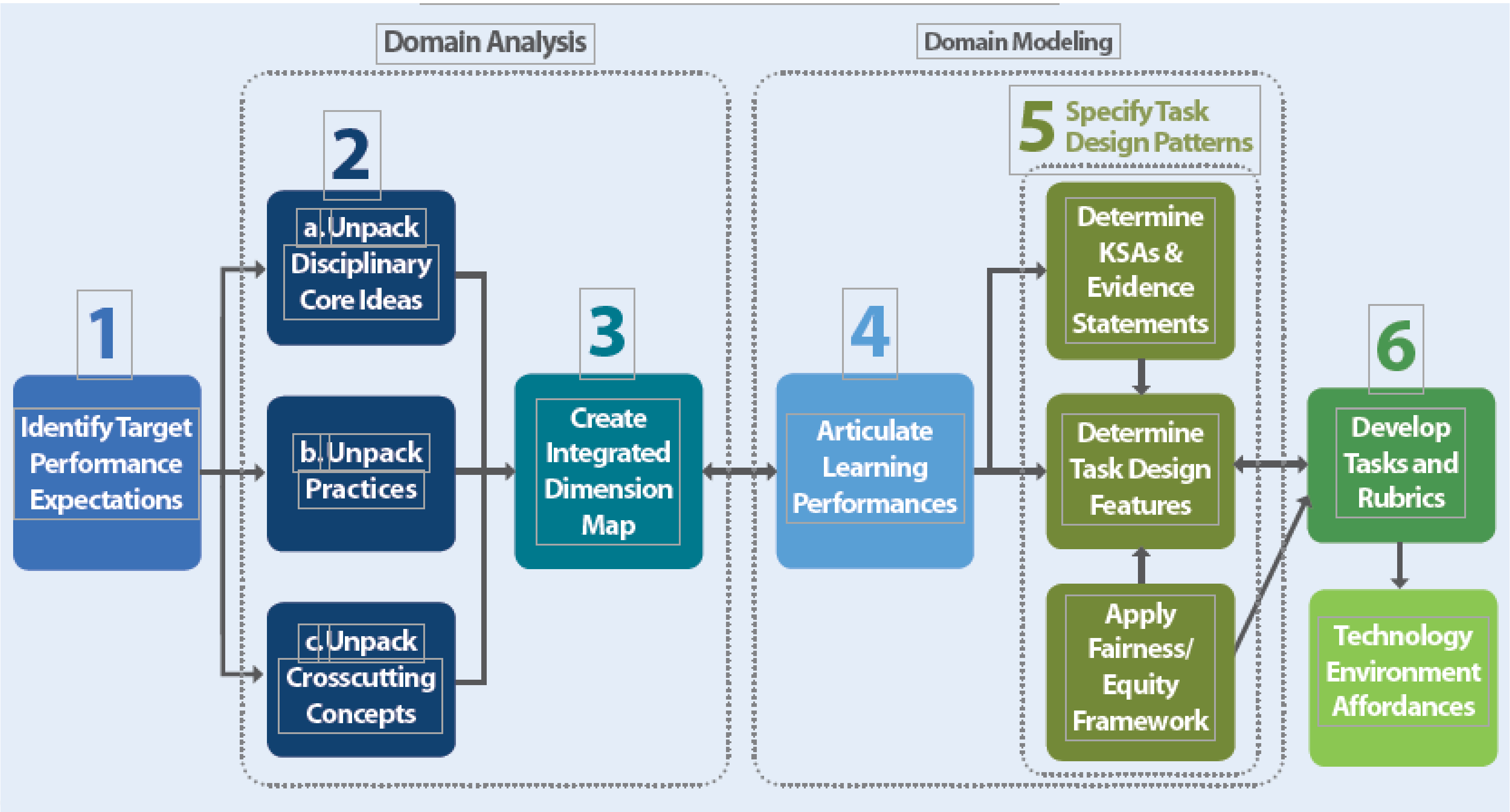


When we have logical and coherent answers to these three questions, we have an *assessment argument*.

How do we Assess *toward* the PEs?

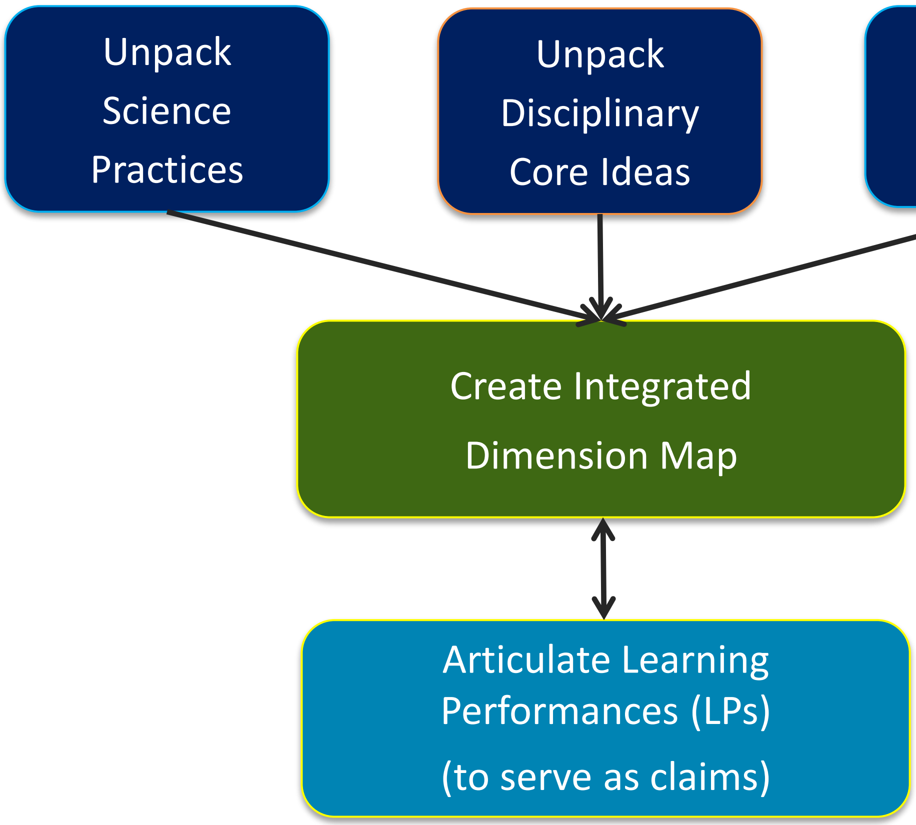
Assess toward Performance Expectations

Principled Task Design – Schematic



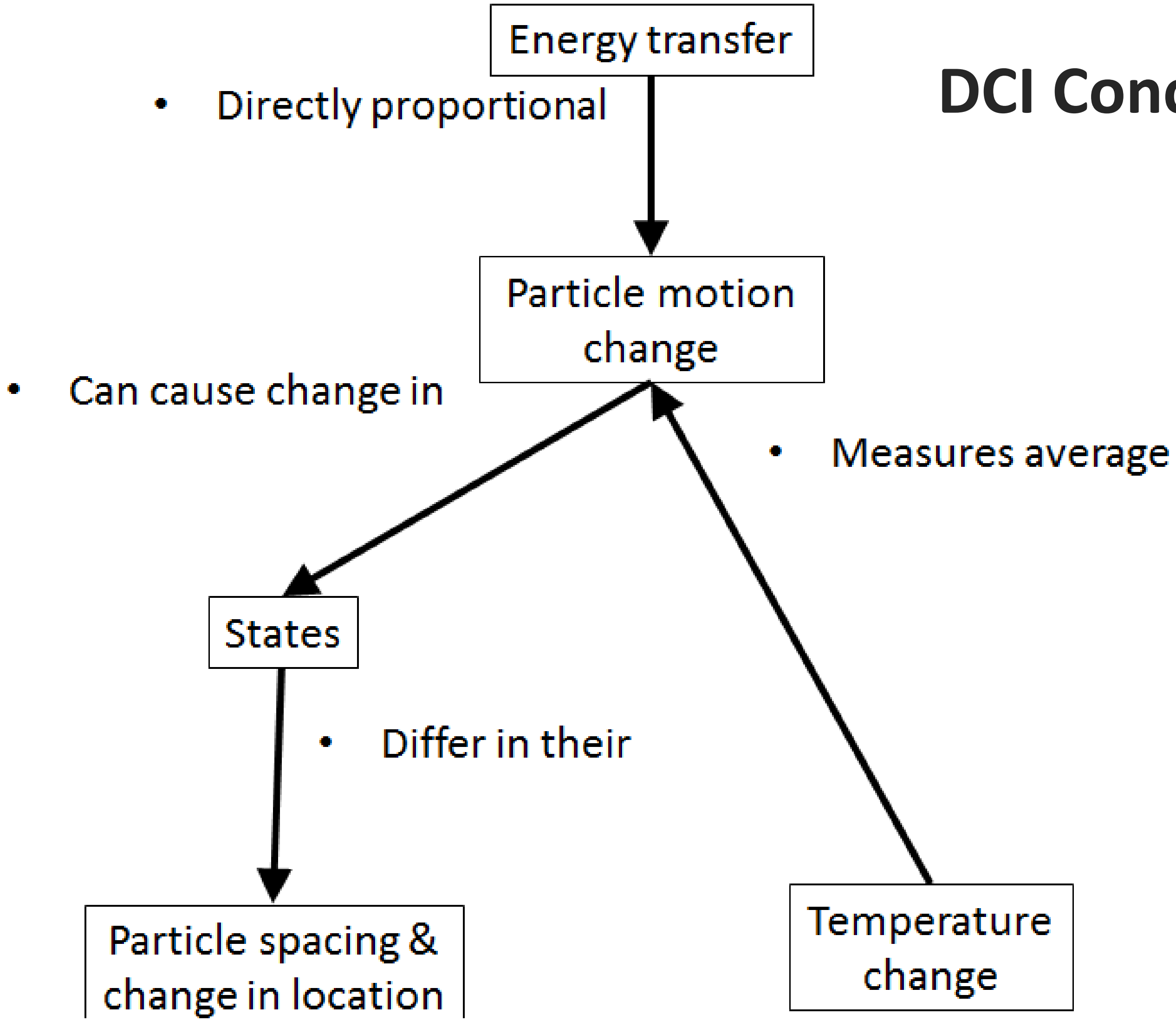
Constructing a Learning Performance

- **Construct a statement or "claim" that you want to make about what a student should be able to do**



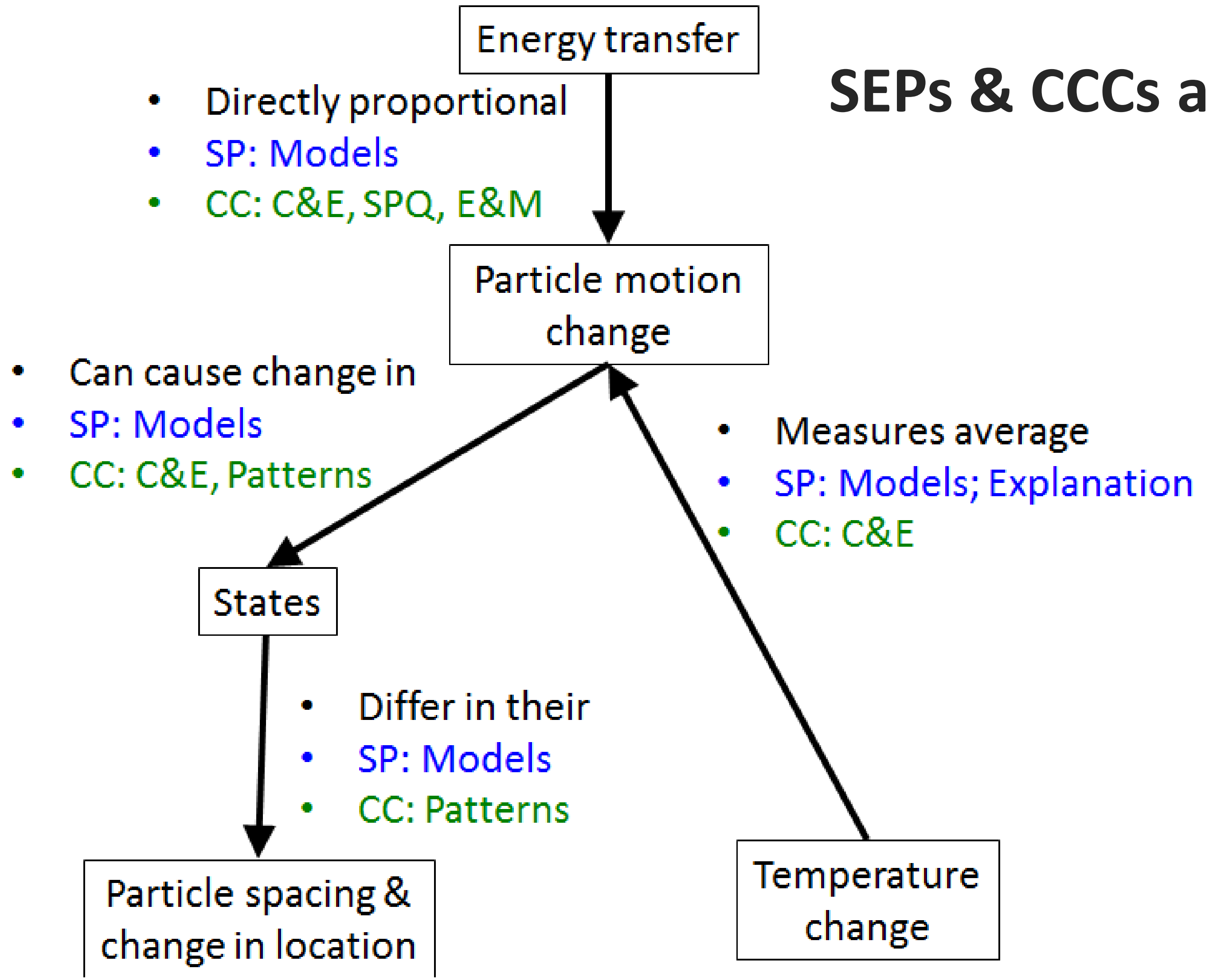
MS-PS1-4. Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.

DCI Concept Map



MS-PS1-4. Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.

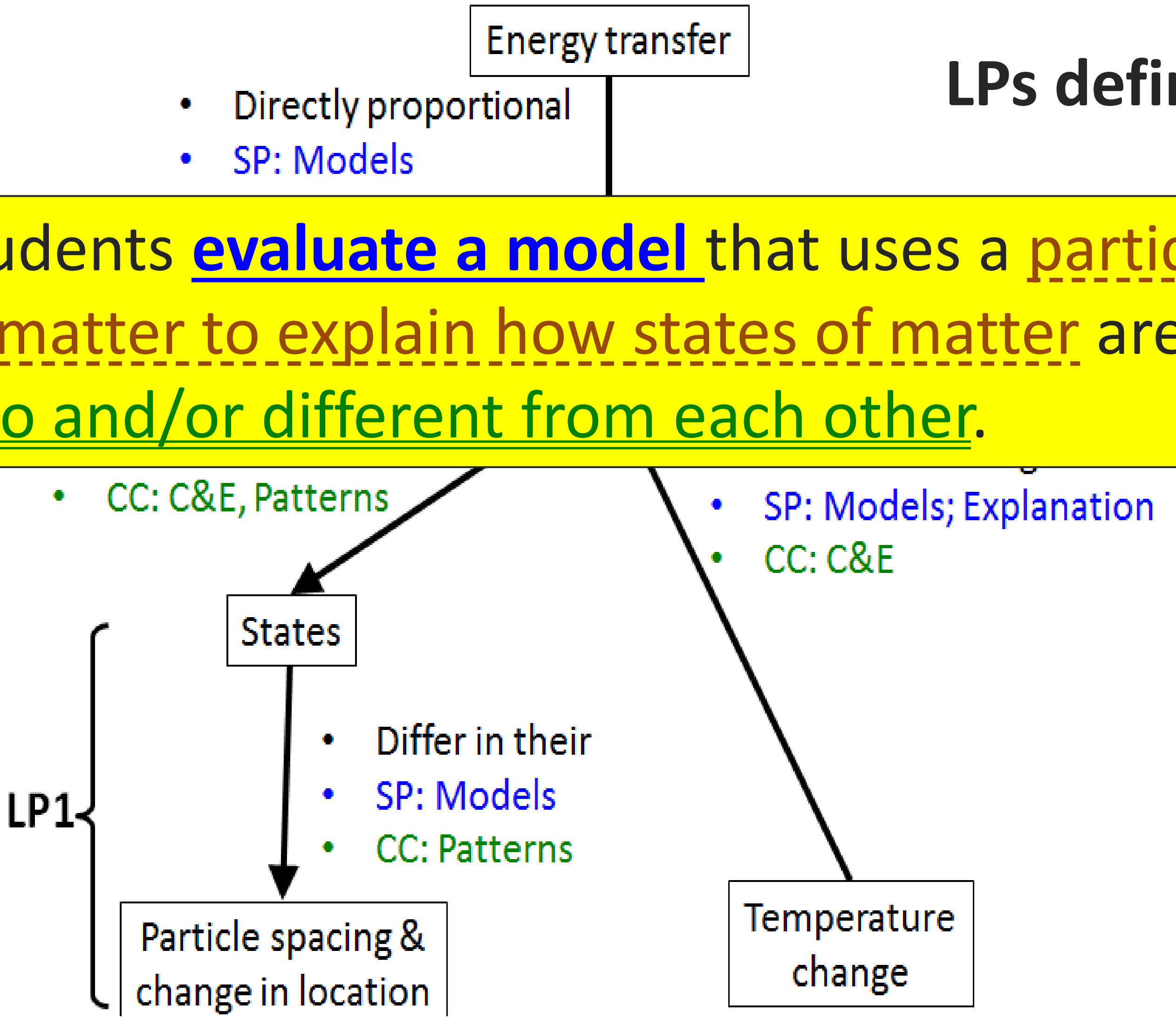
SEPs & CCCs added



MS-PS1-4. Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.

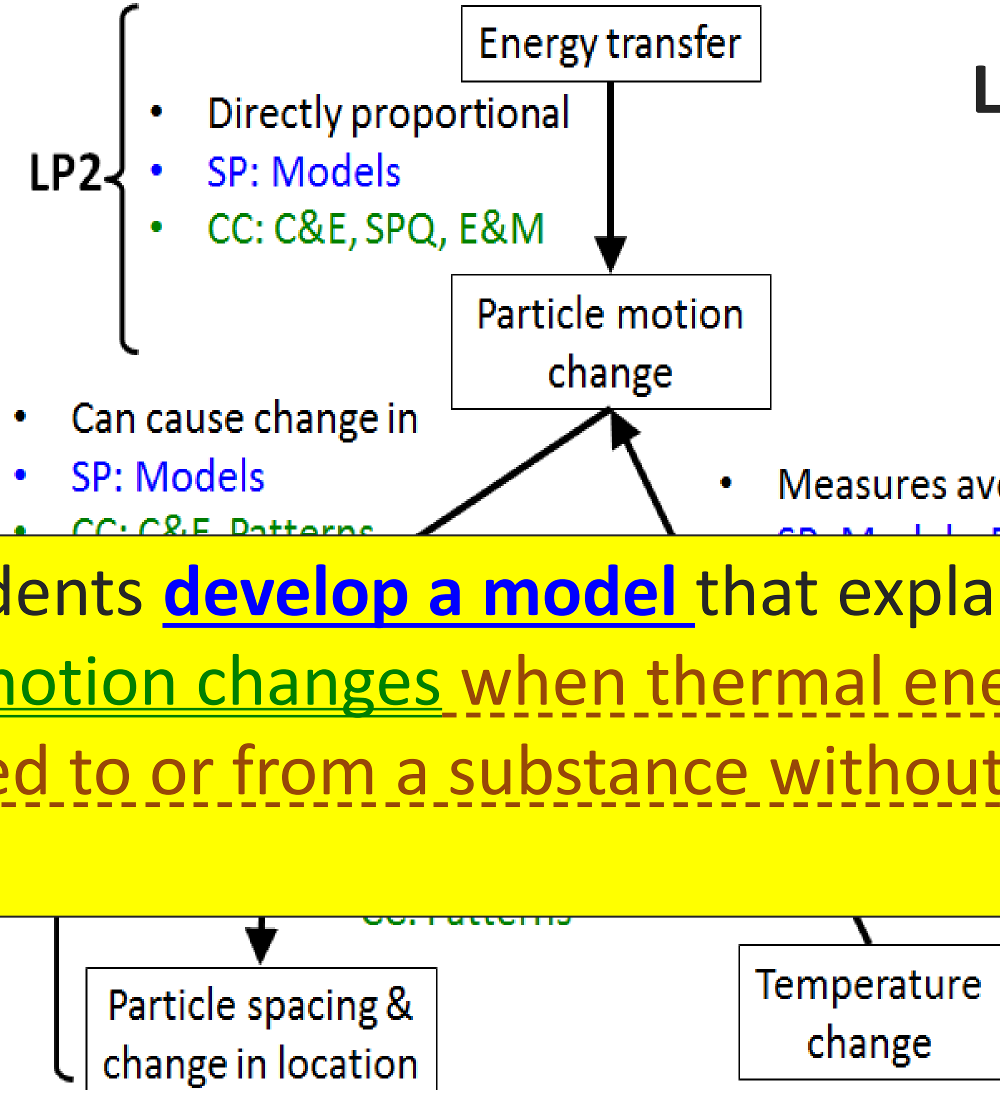
LPs defined

LP 1: Students **evaluate a model** that uses a particle view of matter to explain how states of matter are similar to and/or different from each other.



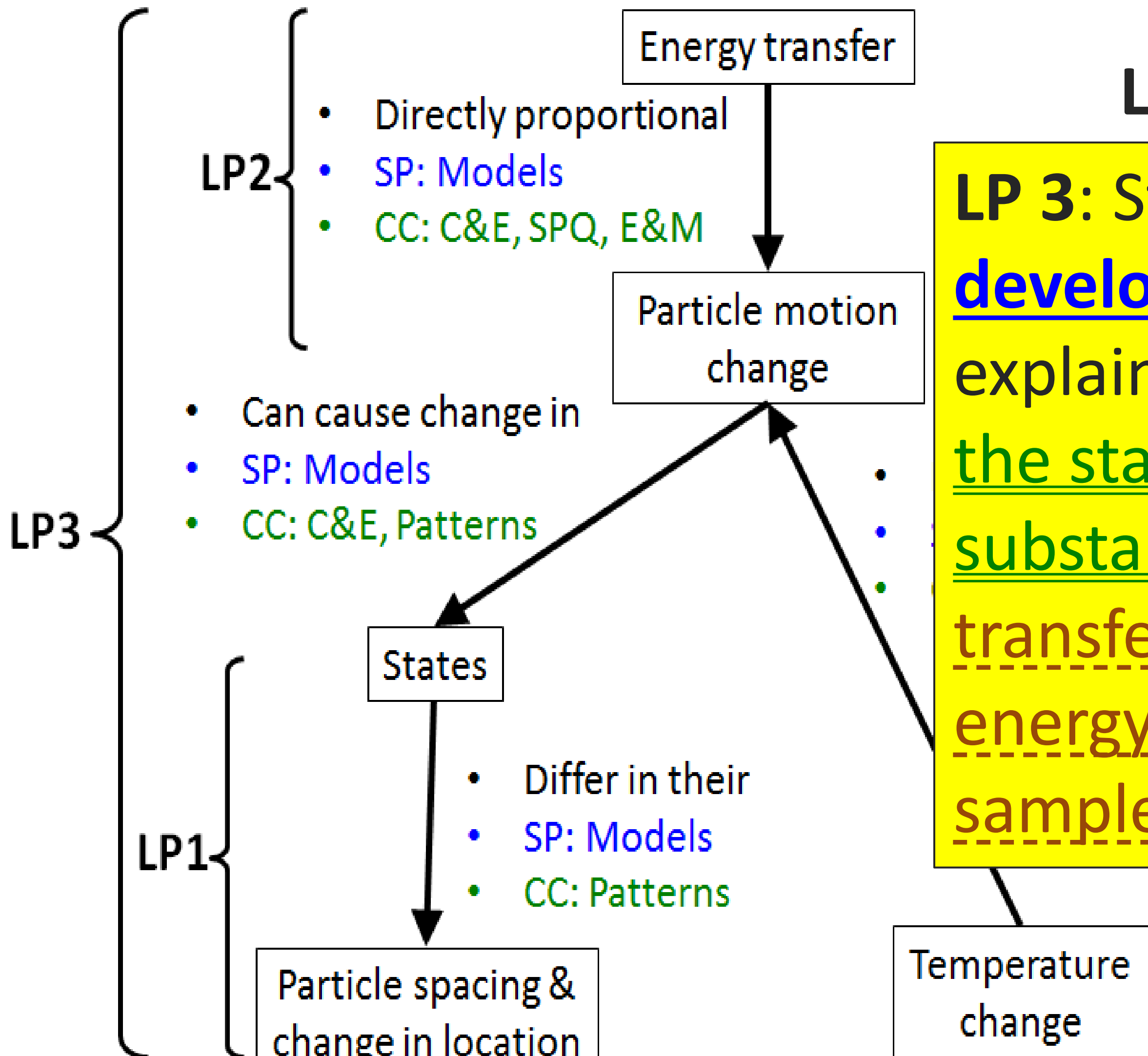
MS-PS1-4. Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.

LPs defined



LP 2: Students **develop a model** that explains how particle motion changes when thermal energy is transferred to or from a substance without changing state.

MS-PS1-4. Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.



LPs defined

LP 3: Students develop a model to explain the change in the state of a substance caused by transferring thermal energy to or from a sample.

Qualities of "Good" Learning Performances

Each **Learning Performance** separately:

- Blends disciplinary core ideas, scientific practices and crosscutting concepts
- Helps to identify an important opportunity that teachers should attend to and assess *before* the end of a unit
- Is assessable in a 5-10 minute task

Collectively the set of all **learning performances**:

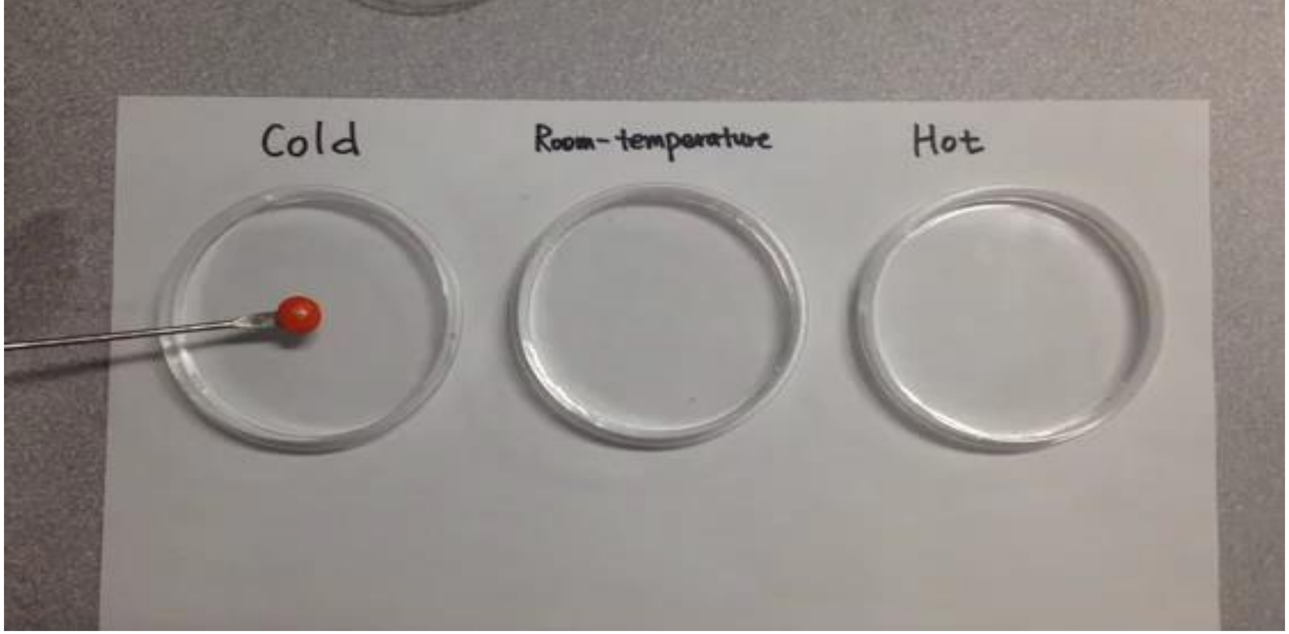
- Identify "what it takes" to make progress toward meeting NGSS performance expectations

Evidence Statement for LP C-02

- Student response should include
- A claim that the properties of the substances indicate that a chemical reaction occurred (or did not occur)
- Evidence supporting the claim (one of the following)
  - Identifies at least one of the available characteristic properties of the substances before and after the process as different (e.g. density, melting point, boiling point, solubility, flammability and odor), in support of a claim that a chemical reaction occurred
  - Identifies all of the available characteristic properties of the substances before and after the process as being the same, in support of a claim that a chemical reaction did not occur
- Reasoning linking the claim and evidence (one of the following)
  - Substances that have the same set of characteristic properties are the same substance
  - Substances that have at least one difference in a characteristic property are not the same substance.
- Reasoning linking the claim and evidence
  - Chemical reactions produce new substances

Specify Task Design Pattern for LP

Construct the Assessment Argument
Specify Learning Performance
Specify Focal Knowledge, Skills, Abilities
Evidence Required to Demonstrate Proficiency
Characteristic Task Features – present in each task
Variable Task Features –varied contexts/difficulty



Shawn had 3 dishes of water at room temperature. She cooled one dish, causing thermal energy to transfer from that dish to the surroundings. She kept the middle dish at room temperature. She transferred thermal energy into the third dish by heating it. Then, Shawn dropped a red-coated chocolate candy into each dish. Watch what happened using the video.

Variable Task Features

- Use of words, graphics, and/or video to present context – **text & video**
- State of matter of substances – **liquid**
- Language demands – **reduced**
- Level of scaffolding to construct model – **yes**

An Example Task for LP C-01

Miranda found four different bottles filled with unknown pure liquids. She measured the mass, volume, and boiling point of these liquids, and also calculated the density of each. The data are displayed in Table 1.

Table 1. Data of four liquids in different bottles.

Sample	Mass	Volume	Density	Boiling Point
1	6.10 g	6.10 cm^3	1.00 g/cm^3	100 C^O
2	5.43 g	6.10 cm^3	0.89 g/cm^3	211 C^O
3	9.38 g	10.20 cm^3	0.92 g/cm^3	298 C^O
4	9.08 g	10.20 cm^3	0.89 g/cm^3	211 C^O

Which information from the data table would you use to determine whether any of the liquids are the same substance? Be sure to tell why.

Based on information in the table, what conclusion can you make about whether any of the liquids are the same? Support your answers with what you know about properties of matter.

Variable Task Features

- Numbers of substances included in the data table.
- State of substances in question (i.e. solid, liquid, or gas state).
- Types and numbers of characteristic and non-characteristic properties included as data.
- Level of scaffolding

SimScientists Assessment Goals

- Design simulation environments that model dynamic science system phenomena "in action"
- Create simulation-based assessments and curriculum supplements to assess and promote NGSS
- Provide evidence of
  - impacts on science learning
  - technical quality
  - feasibility of implementation
  - potential in balanced state science assessment system

SimScientists Projects

Calipers II: Using Science Simulations to Assess Complex Science Learning (NSF)

Foundations of 21st Century Science Assessments (NSF)

Multilevel Assessments of Science Standards (MASS) (IES)

Integrating Science Simulations into Balanced State Science Assessment Systems (OESE)

SimScientists: Interactive Simulation-based Science Learning Environments (IES)

Model Progressions (IES)

SimScientists Assessment Systems (IES)

SimScientists Assessment Systems: Physical Science Links (NSF)

SimScientists Crosscutting Concepts: Progressions in Earth Systems (NSF)

SimScientists Games (NSF)

SimScientists Assessment Suites

(Curriculum-embedded Assessments and Unit Benchmark)

Life Science
- Ecosystems
- Cells
- Human Body Systems
- Genetics
- Evolution

High School
Human Body systems

Physical Science
- Force and Motion
- Atoms and Molecules
- Energy
- Waves

Earth Science
- Geosphere
- Climate

Research Foci

- ***Assessment validity and quality***
  - Alignment with the NGSS an*d* other national and state standards
  - Standards for scientific accuracy/appropriateness, grade-level appropriateness, and item and task quality
  - Psychometric standards for reliability and validity
- ***Classroom use of assessments***
  - Usability
  - Integration into their existing curriculum
  - Value for monitoring learning progress, adjusting instruction, and reporting proficiencies
  - Students' engagement
- ***Policy implications:***
  - Appropriateness as components of a district or state science test
  - Credible components of their state science assessment systems

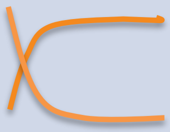
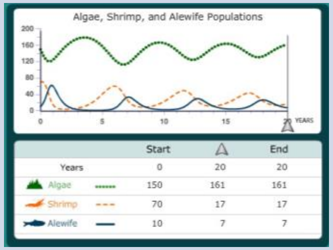
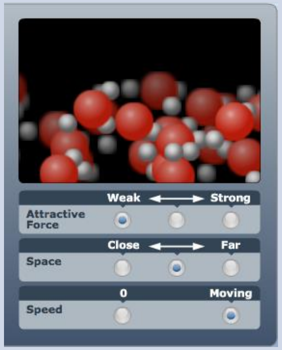
Theory and Research Base

Integrates research on

- *Model-based learning* (Buckley, 2012; Gobert & Buckley, 2000)
    - System Framework-components, interactions, and emergent system behavior
    - The formation, use, evaluation and revision of mental models of complex science systems
- *Evidence-centered assessment design* (Mislevy et al, 2003)
    - A systematic assessment development process that links targets, tasks & data
- *Cognitive science*
    - Guides design and use of representations & interactions in tasks

SimScientists Multidimensional Assessments Exploit Affordances of Technology

- Use a system model framework to develop models of science systems' components, interactions, and dynamic, emergent phenomena
- To understand and apply science knowledge and practices to real world contexts, goals, and problems
- To explain, argue, and critique claims about science phenomena in terms of evidence of the interactions of system components
- To represent and communicate understandings and investigations of dynamic science phenomena

<table>
<tr><th>Content</th><th>Practice</th><th>Integration</th></tr>
<tr>
<td>Populations & changes
</td>
<td>Analyzing and interpreting data</td>
<td>Interpreting patterns of interactions among organisms

</td>
</tr>
<tr>
<td>Forces & collisions
</td>
<td>Developing and using models</td>
<td>Modeling motions and interactions between molecules

</td>
</tr>
</table>

SimScientists 3D Designs

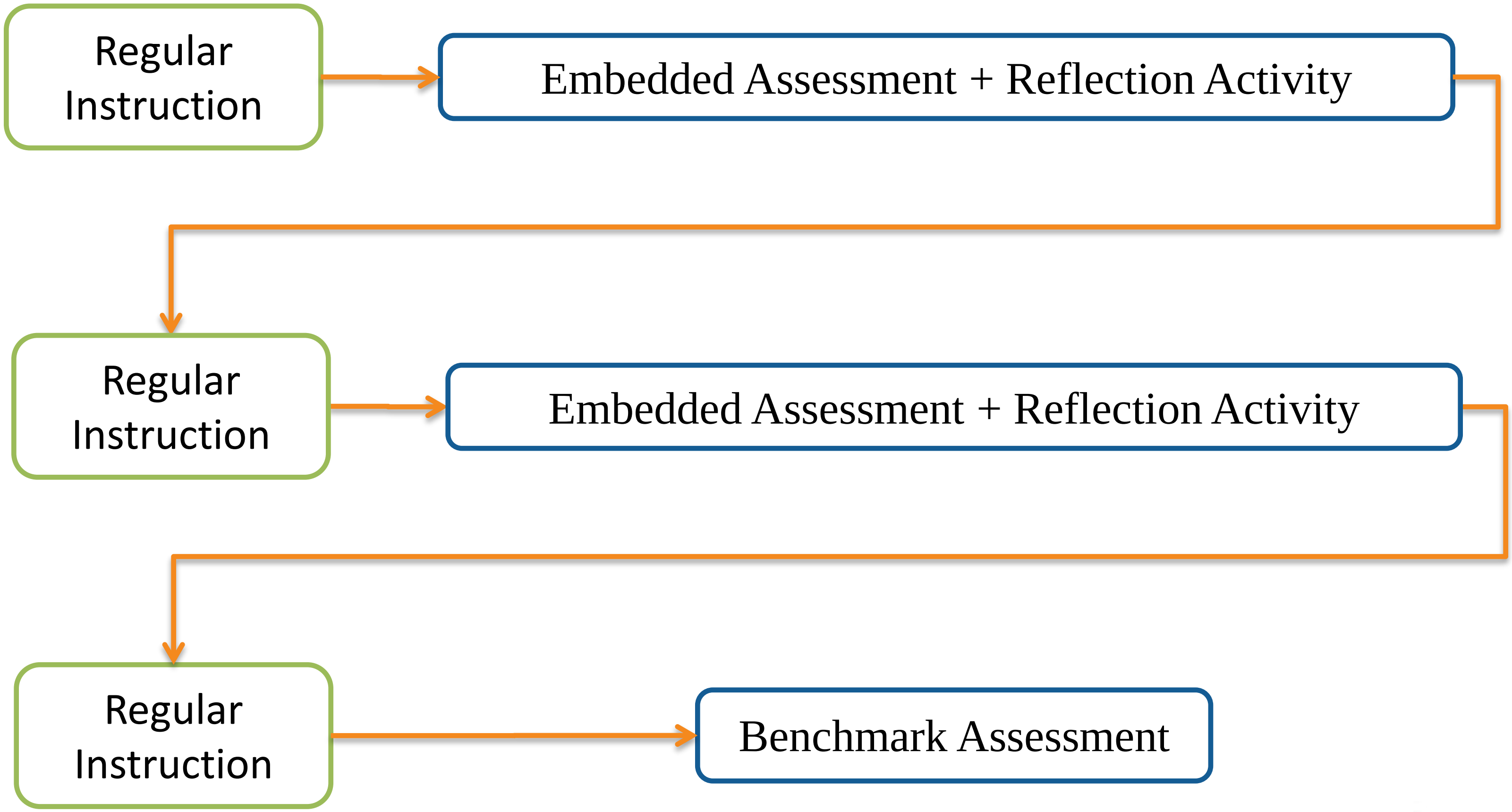
- Simulations both demonstrate and assess integration of core ideas, crosscutting concepts, and practices in investigations
- Use standard terminology of the crosscutting concepts
  - E.g., explicit questions about systems and energy

Task Models

Simulation Environments — for science systems
    physical science, life science, earth science
Embedded Assessments —two to three per topic
    formative
    inserted during unit
    one period
Simulation Benchmark Assessments — one per topic
    summative
    end-of-unit
    one period

SimScientists Assessments

Embedded & Benchmark



Next Generation Science Standards Addressed in Life Science Examples

Cross Cutting concepts

- System and system models
- Energy and matter
- Structure and function

Life science core ideas

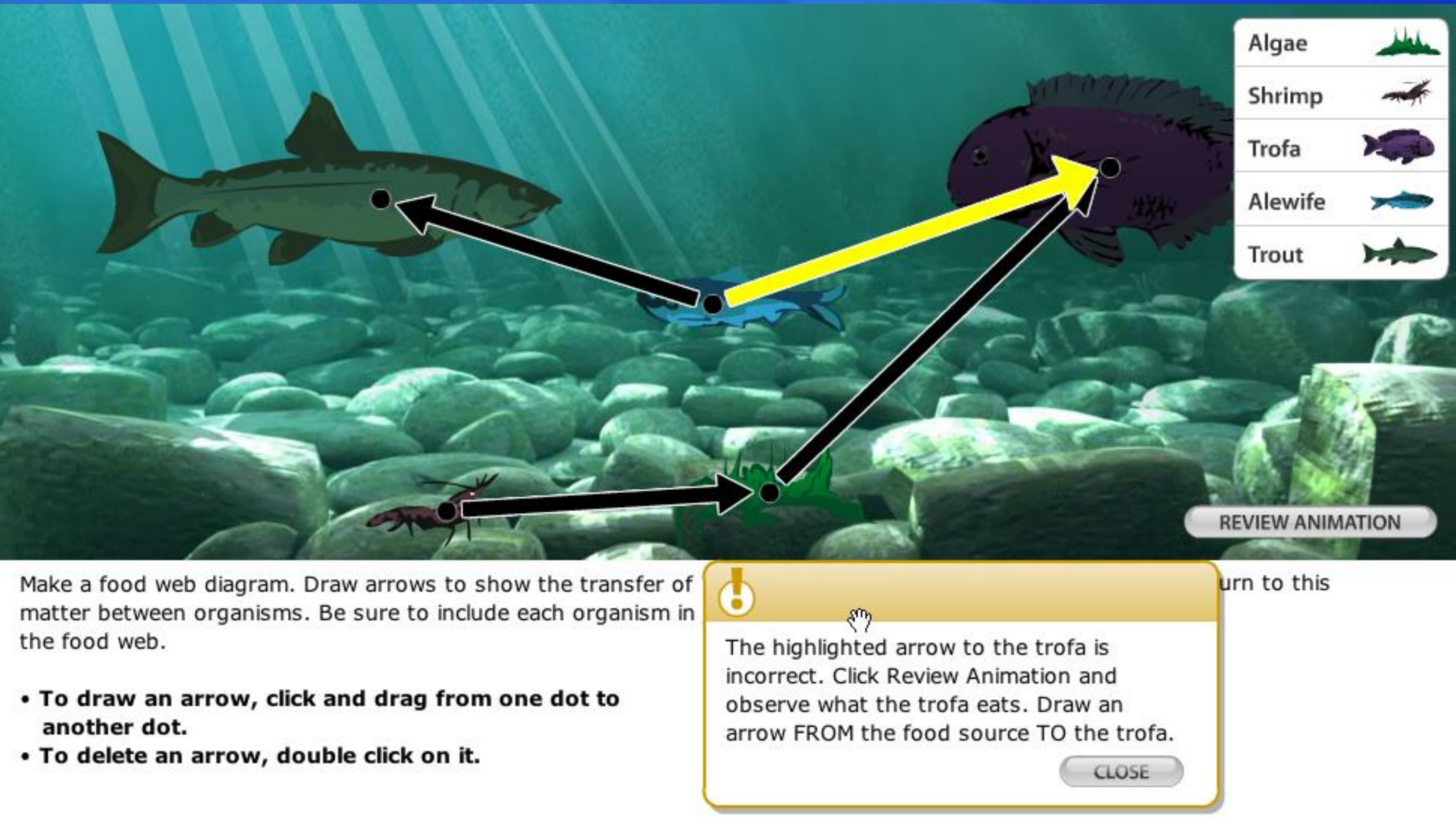
- Ecosystems: Interactions, energy and dynamics
- Human Body Systems: Organ systems work together to maintain a stable internal environment and enable whole body functions

Science practices

- Developing and using models
- Planning, carrying out, analyzing investigations
- Constructing explanations
- Engaging in arguments from evidence

Model Level	Description	Disciplinary Core Ideas	Science Practices
Component 	What are the components of the system and their rules of behavior?	Every ecosystem has a similar pattern of organization with respect to the roles (producers, consumers, and decomposers) that organisms play in the movement of energy and matter through the system. (NGSS: LS2.A — Interdependent relationships in Ecosystems	Analyze and interpret data to provide evidence for phenomena.
Interaction 	How do the individual components interact?	Matter and energy flow through the ecosystem as individual organisms participate in feeding relationships within an ecosystem. (NGSS: LS2.B — Cycles of Matter and Energy Transfer in Ecosystems)	Develop a model to describe phenomena. Analyze and interpret data
Emergent 	What is the overall behavior or property of the system that results from many interactions following specific rules?	Interactions among organisms and among organisms and the ecosystem's nonliving features cause the populations of the different organisms to change over time. (NGSS: LS2.C — Ecosystem Dynamics, Functioning, and Resilience)	Use a model to plan and carry out investigations. Analyze and interpret data to provide evidence.

SimScientists Curriculum Embedded Assessments



Affordances of Simulation-based Assessments for Formative Purposes

- Students can be engaged in standards-based, cognitively-principled assessment tasks designed to elicit evidence of understanding of science system models and active application of science practices.
- Assessment tasks can represent often invisible dynamic science system phenomena “in action” and
- The simulations allow students to conduct multiple, iterative investigations of how science phenomena change under different conditions.
- As students engage in investigations, they receive immediate, individualized feedback, and adaptive, additional instruction.

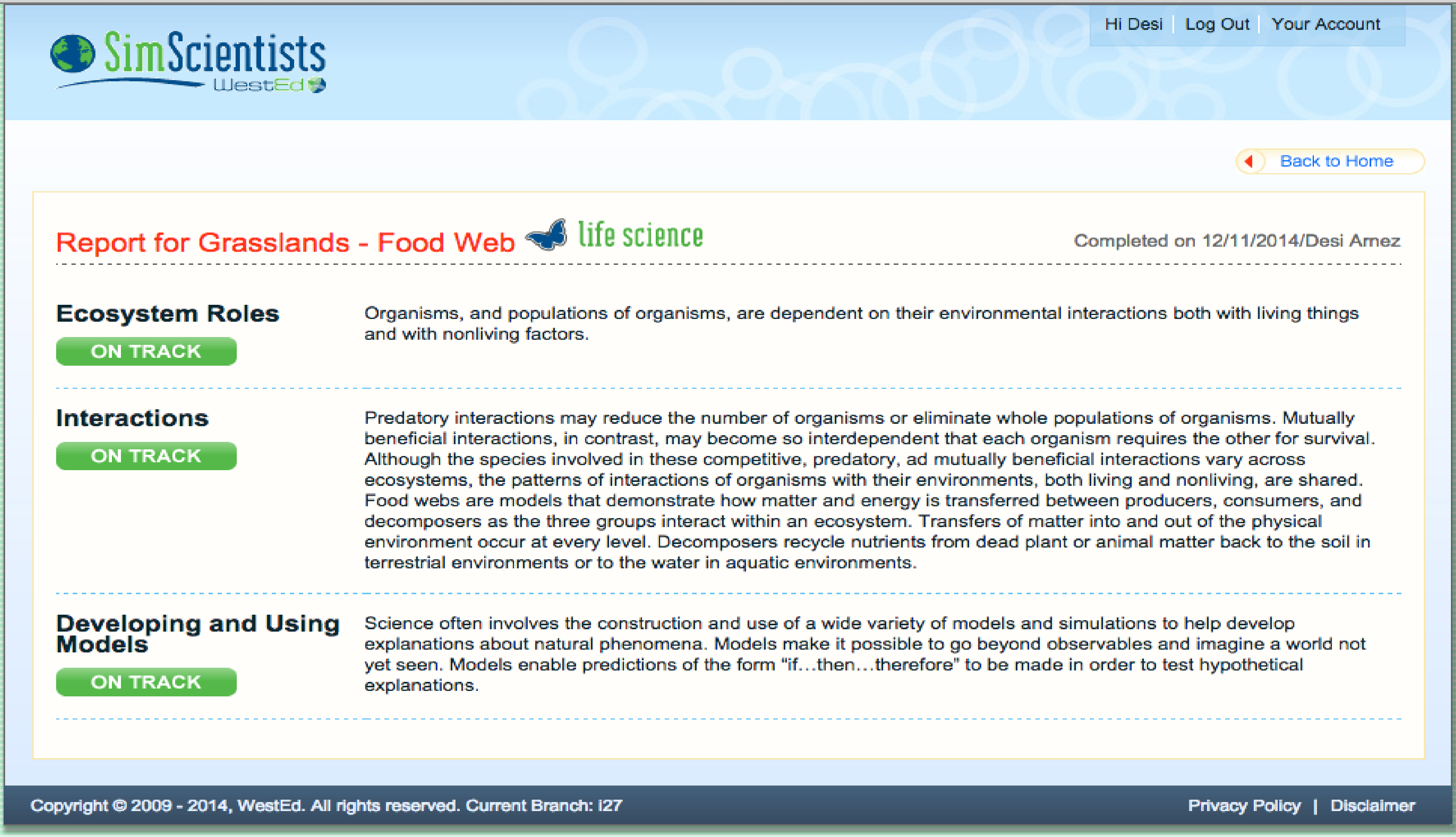
Affordances of Simulation-based Assessments for Formative Purposes (2)

- The simulation software provides instant reports of individual and class performance on multiple core ideas and science practices-for students and teachers
- The teacher can monitor individual student and class engagement and performance during the simulation—online and circulating
- Formative use of simulation-generated progress reports
  - The learning management system can recommend assignment of students to teams based on the embedded-assessment results

Evidence Model

- Auto-scored selected responses, arrows, etc
- Constructed responses
  - Rubrics for teachers and students
- Benchmark assessments
  - Score reports by standard/target
  - Bayesian networks
- Embedded assessments
  - State-based progress levels

Sample Progress Reports to Students Ecosystems



Individual Progress Report: Populations

Populations	ON TRACK	Ecosystems are dynamic in nature; their characteristics can vary over time. Disruptions to any physical or biological component of an ecosystem can lead to shifts in all its populations.
Developing and Using Models	ON TRACK	Science often involves the construction and use of a wide variety of models and simulations to help develop explanations about natural phenomena. Models make it possible to go beyond observables and imagine a world not yet seen. Models enable predictions of the form “if…then…therefore” to be made in order to test hypothetical explanations.
Planning and Carrying Out Investigations	NEEDS HELP	Scientific investigations may be conducted in the field or the laboratory. A major practice of scientists is planning and carrying out a systematic investigation, which required the identification of what is to be recorded and, if applicable, what are to be treated as the dependent and independent variables (control of variables). Observations and data collected from such work are used to test existing theories and explanations or to revise and develop new ones.
Analyzing and Interpreting Data	ON TRACK	Scientific investigations produce data that must be analyzed in order to derive meaning. Because data usually do not speak for themselves, scientists use a range of tools- including tabulation, graphical interpretation, visualization, and statistical analysis- to identify the significant features and patterns in the data. Sources of error are identified and the degree of certainty calculated. Modern technology makes the collection of large data sets much easier, thus providing many secondary sources for analysis.
Constructing Explanations	NEEDS HELP	The goal of science is the construction of theories that can provide explanatory accounts of features of the world. A theory becomes accepted when it has been shown to be superior to other explanations, in the breadth of phenomena it accounts for, and its explanatory coherence and parsimony. Scientific explanations are explicit applications of theory to a specific situation or phenomenon, perhaps with the intermediary of a theory-based model for the system under study. The goal for students is to construct logically coherent explanations of phenomena that incorporate their current understanding of science, or a model that represents it, and are consistent with the available evidence.

Discussion

- Both projects exemplify multi-dimensional science learning and assessment.
  - What are the advantages and limitations of multi-dimensional assessment?
  - What assumptions are included in inferences about student proficiency drawn from evidence elicited in multi-dimensional tasks?
- Both projects share approaches to assessment design, incorporating Evidence-Centered Design and specifying design patterns for eliciting evidence of student proficiency.
  - What do members of each project team see reflected of their approaches in the work exhibited by the other team?
- Each project starts from a different point, either by developing learning performances that synthesize multiple dimensions of the Framework for K-12 Science Education or by directly targeting existing performance expectations in the Next Generation Science Standards.
  - Within each project, what is the rationale for the focus on one or the other, and how does the educational context for each project inform this decision?
  - What do educators, researchers, and test developers need to know about the role of each document in the broader science education landscape?

Summary & Closing Discussion

For More information about our work

NGSA Project

Website: **www.nextgenscienceassessment.org**

Contacts: ***Joseph Krajcik*** **– krajcik@msu.edu (Twitter: @krajcikjoe)**

Christopher Harris **– christopher.harris@sri.com**

Lou DiBello **–ldibello@uic.edu**

SimScientists Project

Website: **http://simscientists.org**

Contacts: ***Edys Quellmalz–*** **equellm@wested.org**

Matt Silberglitt **– msilber@wested.org**

These projects are funded in part by the National Science Foundation, grant numbers 1316903, 1316908, 1316874, and XX) . Any opinions, findings, and conclusions or recommendations at this session or in these materials are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.