SimScientists: Using Simulations to Assess Next Generation Science Learning

Edys Quellmalz and Matt Silberglitt

WestEd

Presented to the Workshop on Developing Assessments to Meet the Goals of the 2012 Framework for K-12 Science Education. National Academy of Sciences, WDC September 13, 2012

This material is based upon work supported by the National Science Foundation under Grant # 0733345, awarded to WestEd and by the US Department of Education under OESE Grant# 09-2713-126 awarded to the Nevada Department of Education, and under IES Grant # R305A080225, awarded to WestEd. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or US Department of Education.

SimScientists Assessment Team

- Barbara Buckley
- Mark Loveland
- Daniel Brenner
- Jodi Davenport
- Kevin Jordan
- Mike Timms*

**Currently at the Australian Council for Education Research (ACER)*

SimScientists WestEd

Assessment R&D Projects

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

Multilevel Assessments of Science Standards (MASS) (IES)

Integrating Science Simulations into State Science Assessment Systems Enhanced Assessment Grant (EAG) (USDOE:OESE)

SimScientists Assessment Systems (IES)

SimScientists Assessment Physical Science Links (NSF)

Foundations of 21st Century Assessment (NSF)

SimScientists Assessment Purposes

- To assess use of cross-cutting concepts, core ideas, and inquiry practices to study science systems
- To develop and document technical quality, feasibility, and instructional utility of simulation-based assessments
- To validate use of simulations in multilevel, balanced science assessment systems

SimScientists WestEd

Relevance to 2012 Framework for Science Education

- Cross-Cutting Concepts
  - Systems and Models
  - Energy and Matter
- Inquiry practices for using knowledge in significant, authentic tasks
- Core ideas

SimScientists Assessment Development

- Learning science research
  - Model-based learning
  - Science inquiry
- Multimedia research
  - Types of representations
  - Forms of response expression
- Measurement research
  - IRT, Bayes Nets, and Markov Decision Processes
- Evidence-centered assessment design

Technology Affordances

- Animations of dynamic system phenomena
  - Can observe and review
- Simulation-based investigations
  - Iterative design
  - Virtual data collection
  - Conducting and saving multiple trials
  - Multimodal information and data displays
- Multiple, overlapping, simultaneous representations
- Scientific “tools of the trade”
  - Simulations, graphs, tables, zoom, drawing, highlighting
- Immediate, contingent feedback and hints
- Bayes Nets within simulations to assess proficiencies

Principled Assessment Design (Mislevy et al., 2003)

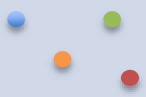
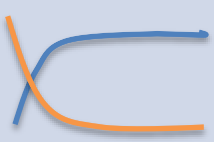
Student Model ⟷ Evidence Model ⟷ Task Model

Student Model	Evidence Model	Task Model
What complex of knowledge, skills, or other attributes should be assessed?	What behaviors or performances should reveal the relevant knowledge and skills described in the student model?	What tasks or situations should elicit the behaviors or performances described in the evidence model?

SimScientists Student Model

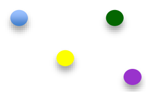
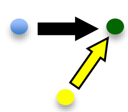
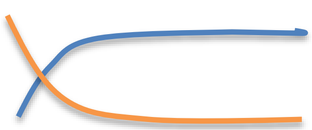
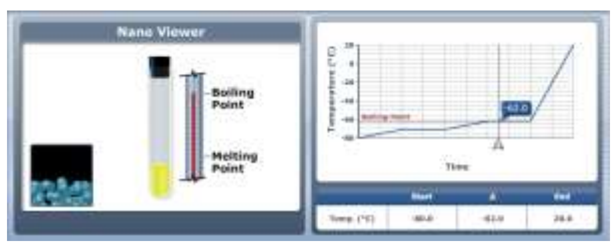
- Science System Model Levels to frame core ideas
  - Components
  - Interactions
  - Emergent system behavior
- Science investigations practices
    - E.g., Plan, conduct, analyze, interpret, communicate

System Model for Middle School Ecosystems

Model Level	Descriptions	Concept Targets	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)	Analyzing and Interpreting Data
Interaction 	How do the 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)	Developing & Using Models; Analyzing and Interpreting 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—Ecosystems Dynamics, Functioning and Resilience)	Planning and Carrying Out Investigations; Analyzing and Interpreting Data

SimScientists WestEd

Atoms & Molecules Target System Model

Component	Atoms and Molecules	Skill
	Nitrogen · Water Vapor · Argon	Observe
Interaction	**Speed – Spacing – Collisions**	**Skill**
	 	Analyze
Emergent	**Boiling & Melting Point – States of Matter**	**Skill**
		Measure & Investigate

SimScientists WestEd

Task Model Features

- Overarching goal or problem
- Model level progressions for integrating
  - Components, interactions, emergent behavior
  - Inquiry practices
    - Observing components,
    - Inferring roles and interactions
    - Predicting, observing, explaining
    - Designing, conducting, analyzing simulation-based investigations
    - Critiquing designs of others

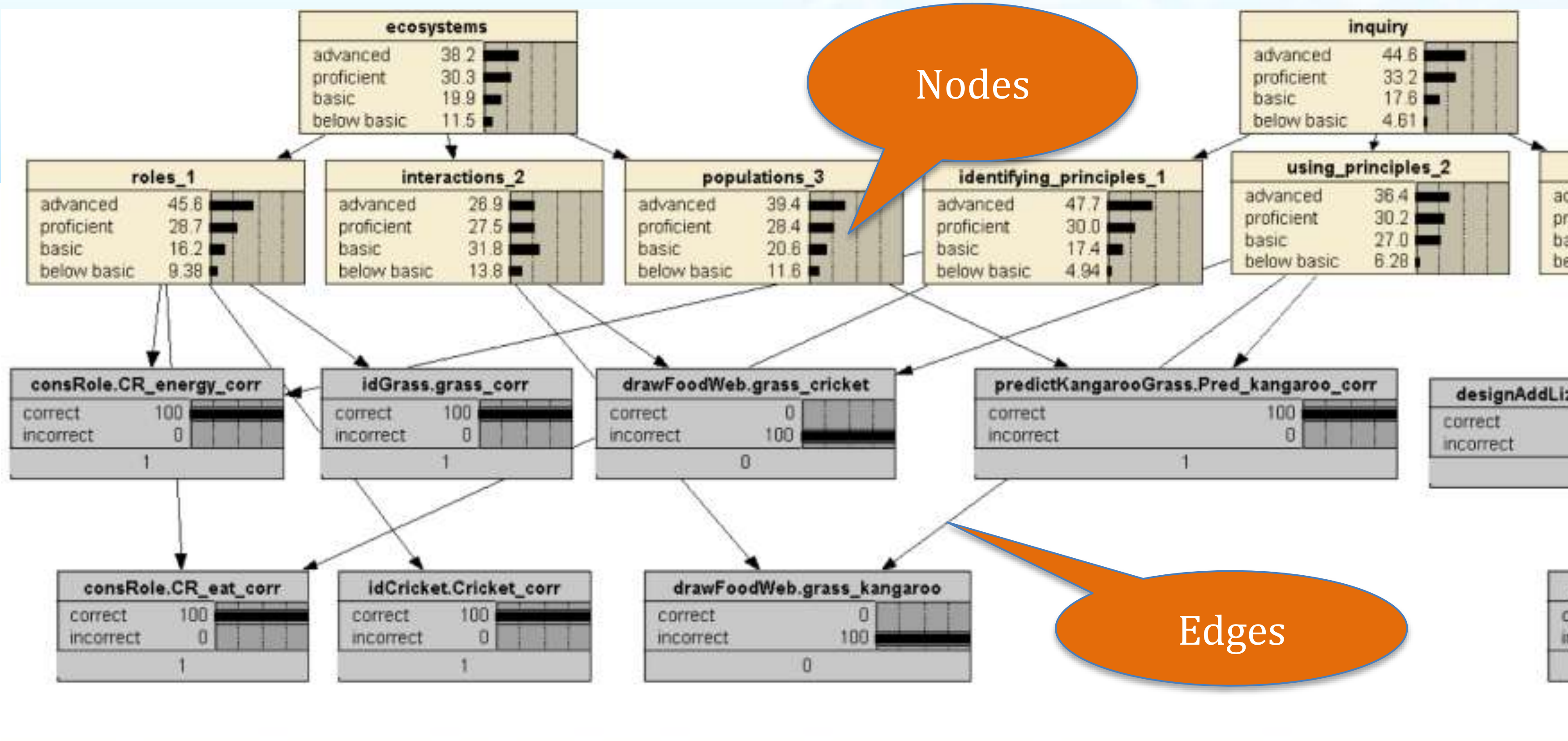
Task Model Features (cont'd)

- Curriculum-embedded assessments (for formative use)
  - Feedback and graduated coaching
  - Self-assessment of explanations
  - Progress reports
  - Reflection activities off-line to adjust instruction
- Benchmark assessments (for summative purposes)
  - Parallel tasks to assess same content and inquiry targets
  - No feedback or coaching
  - Bayes Net to report proficiency levels
- Signature Tasks (for summative purposes)

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
  - Algorithm-based progress levels

SimScientists WestEd



Fragment of a Bayes Net From the SimScientists Ecosystems Benchmark Assessment

Note: the conditional probabilities associated with the edges are not visible in this view

Components of the SimScientists Classroom Assessments

Embedded in Classroom Instruction

Embedded Formative Assessments and Reflection Activities (2 or 3)

Online module with feedback and coaching → Progress Report → Follow up Classroom Reflection Activity

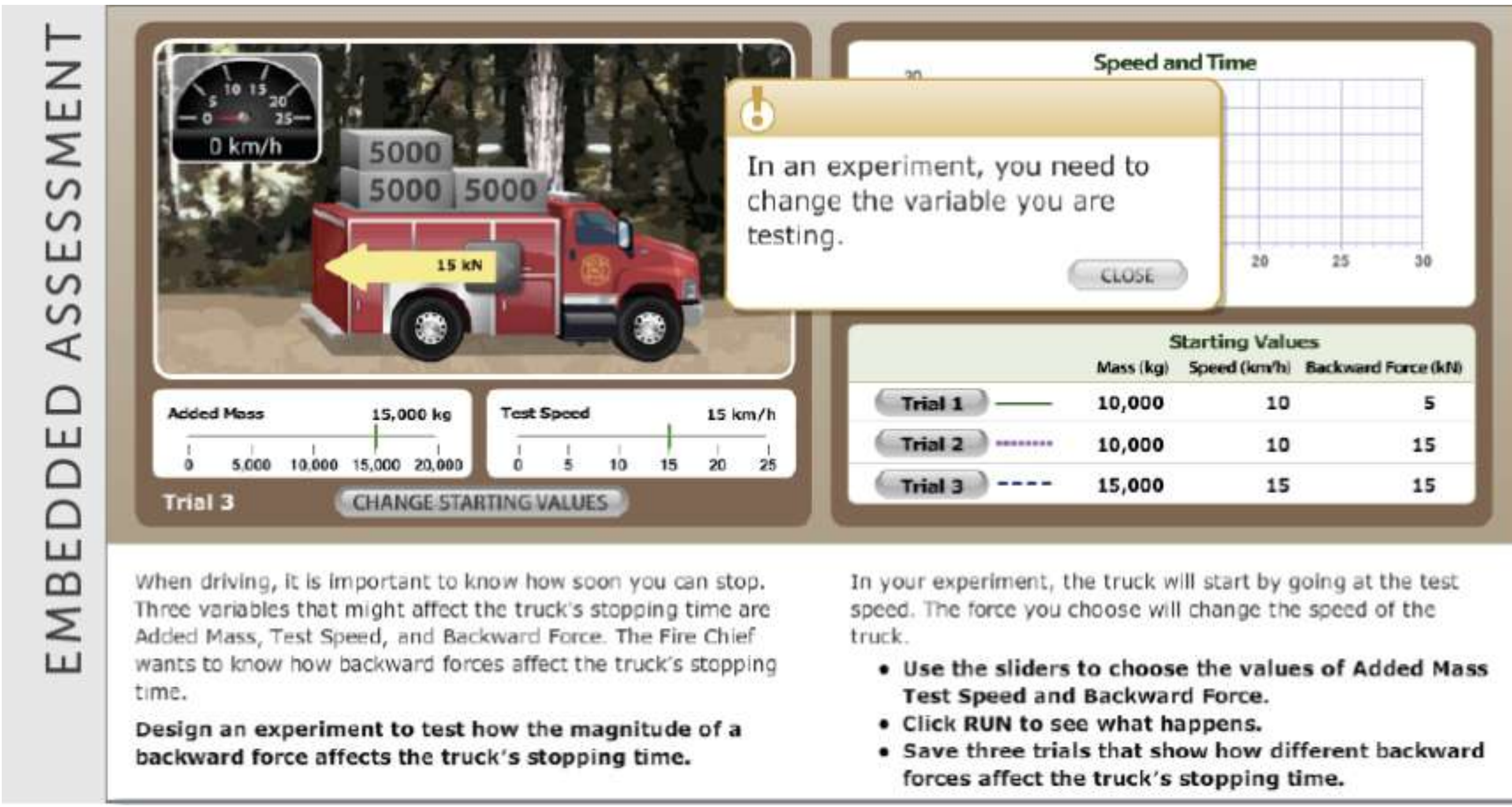
Benchmark Summative Unit Assessments

Online assessment without feedback → Teacher scores constructed responses → LMS → Proficiency Report

SimScientists WestEd

Examples of the Assessment Modules

Embedded Assessment



Embedded Assessment Progress Report to Student

Report for Mountain Lake - Predator Prey

life science

Completed on 03/23/2010/Sara

Populations ON TRACK	Interactions between organisms and between organisms and the ecosystem's nonliving features cause the populations of the different organisms to change over time.
Conduct ON TRACK	Conducting investigations involves carrying out scientific investigations using appropriate tools and techniques.
Identify NEEDS HELP	Identifying Science Principles focuses on students' ability to recognize, recall, define, relate, and represent basic science principles. The practices assessed in this category draw on declarative knowledge or "knowing that."
Design NEEDS HELP	Designing investigations involves asking questions, planning investigations and evaluating experimental design.
Analyze PROGRESSING	Identifying patterns involves summarizing patterns in data, analyzing which data are relevant and drawing conclusions by relating patterns in data to theoretical models.

SimScientists WestEd

Embedded Assessment Progress Report for Teacher: Class Summary

Summary Report: Mountain Lake - Food Web

Try it | Detailed Report

ASSESSMENT: Mountain Lake - Food Web

CLASS: Period 7

Go!

Content	NH Needs Help	P Making Progress	OT On Track
Roles	12 (46%)	4 (15%)	10 (38%)
Interactions	15 (58%)	4 (15%)	7 (27%)
Inquiry			
Identifying	15 (58%)	5 (19%)	6 (23%)
Using	10 (38%)	5 (19%)	11 (42%)

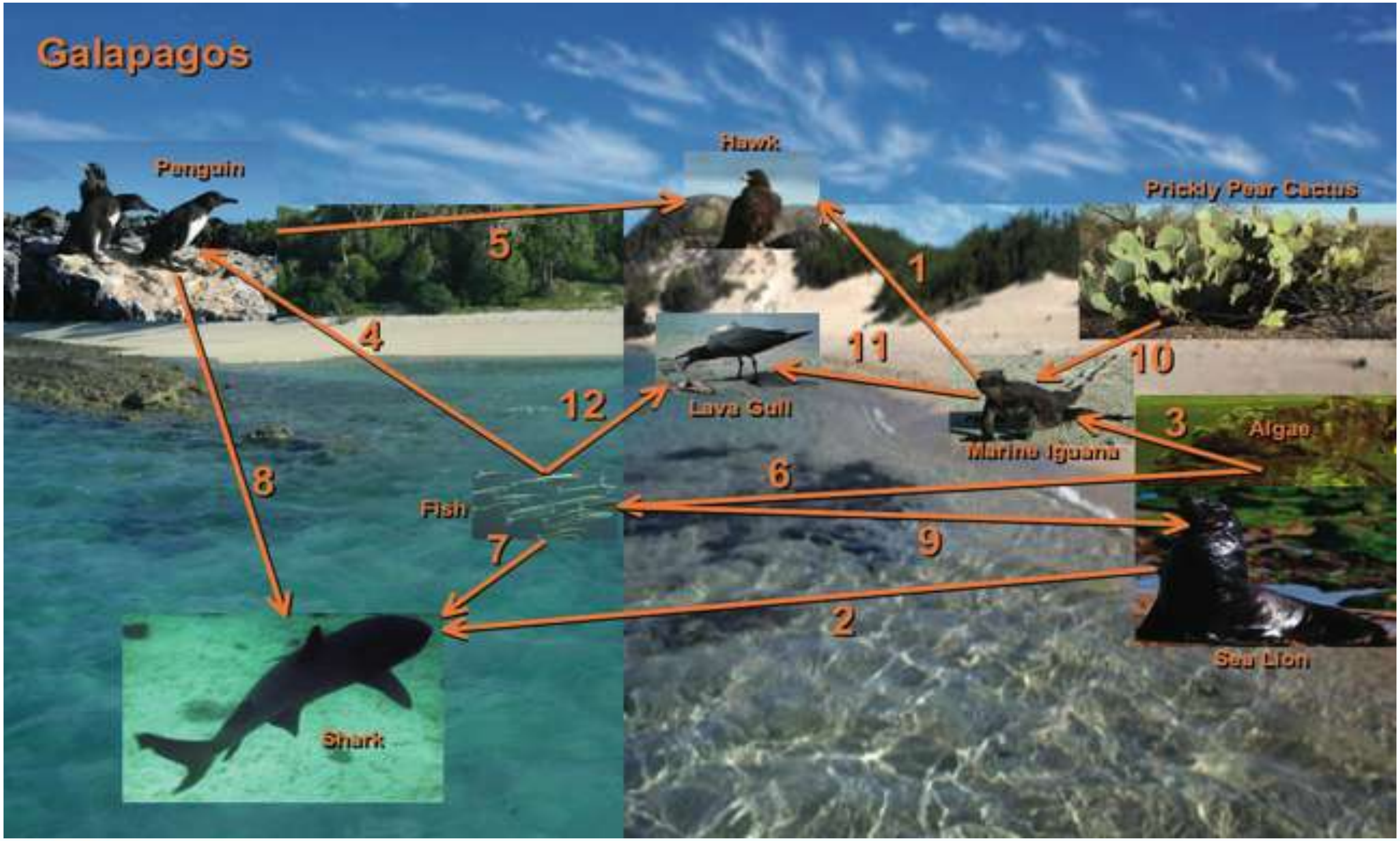
NH = needs help P = making progress OT = on track

SimScientists WestEd

Classroom Reflection Activity

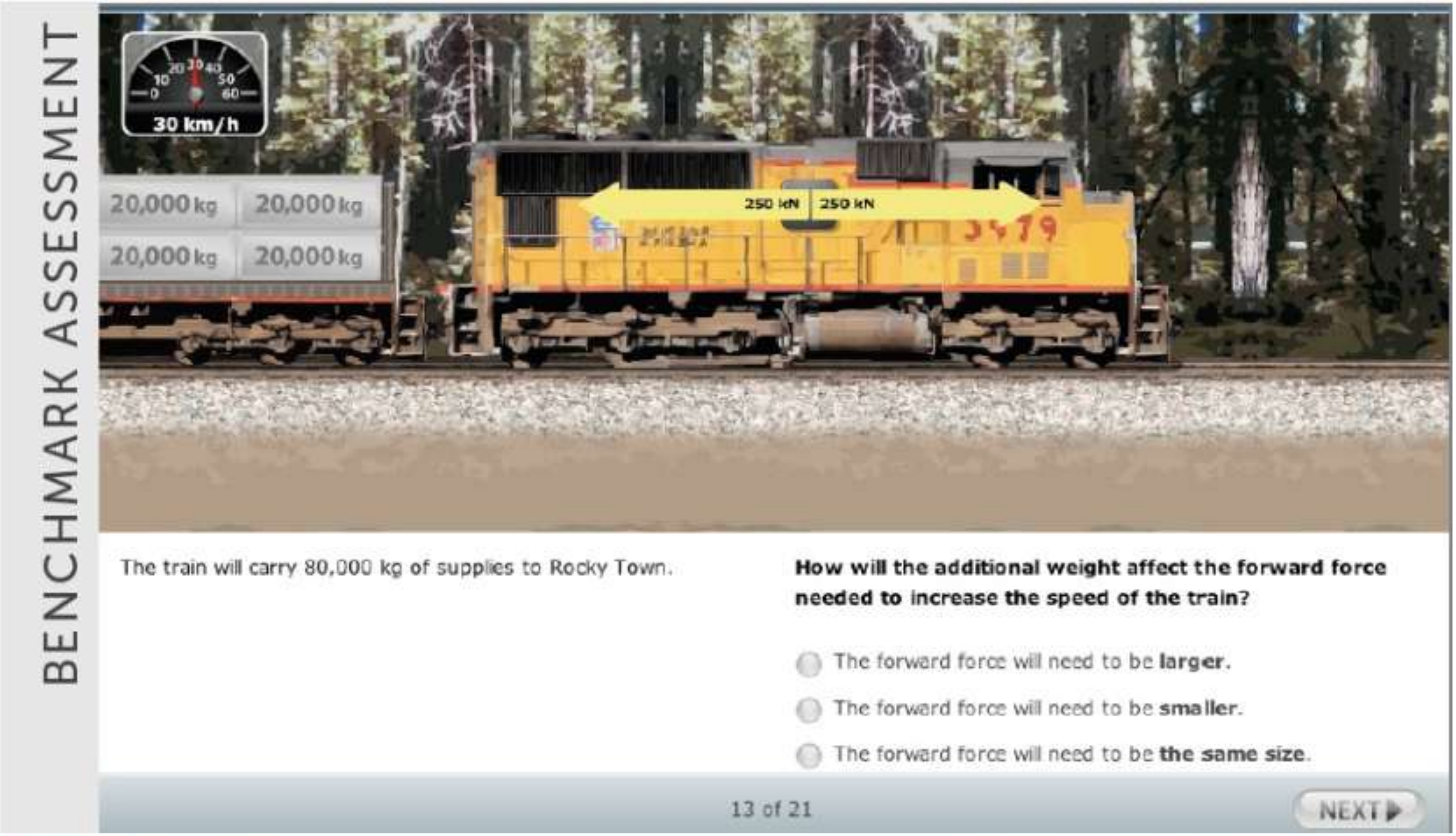
- Formative use of assessment results
- Transfer to different, more complex system
- Collaborative jigsaw structure
- Communication

Transfer to new, more complex ecosystem



SimScientists WestEd

Benchmark Assessment



SimScientists WestEd

Summary Benchmark report

Content		BB Below Basic	B Basic	P Proficient	A Advanced
Roles	BB	3 (100%)	0 (0%)	0 (0%)	0 (0%)
Interactions	BB	3 (100%)	0 (0%)	0 (0%)	0 (0%)
Populations	A	0 (0%)	0 (0%)	0 (0%)	3 (100%)
Inquiry					
Identify	BB A	1 (33%)	0 (0%)	0 (0%)	2 (67%)
Use	BB B	1 (33%)	2 (67%)	0 (0%)	0 (0%)
Design	BB B	1 (33%)	2 (67%)	0 (0%)	0 (0%)
		0 (0%)	3 (100%)	0 (0%)	0 (0%)

Interactions: Matter and energy flow through the ecosystem as individual organisms interact with each other. Food web diagrams indicate the feeding relationships among organisms in an ecosystem. All ecosystems have a flow of energy from a nonliving source, to producers, to consumers.

Detailed Report by Student and Target

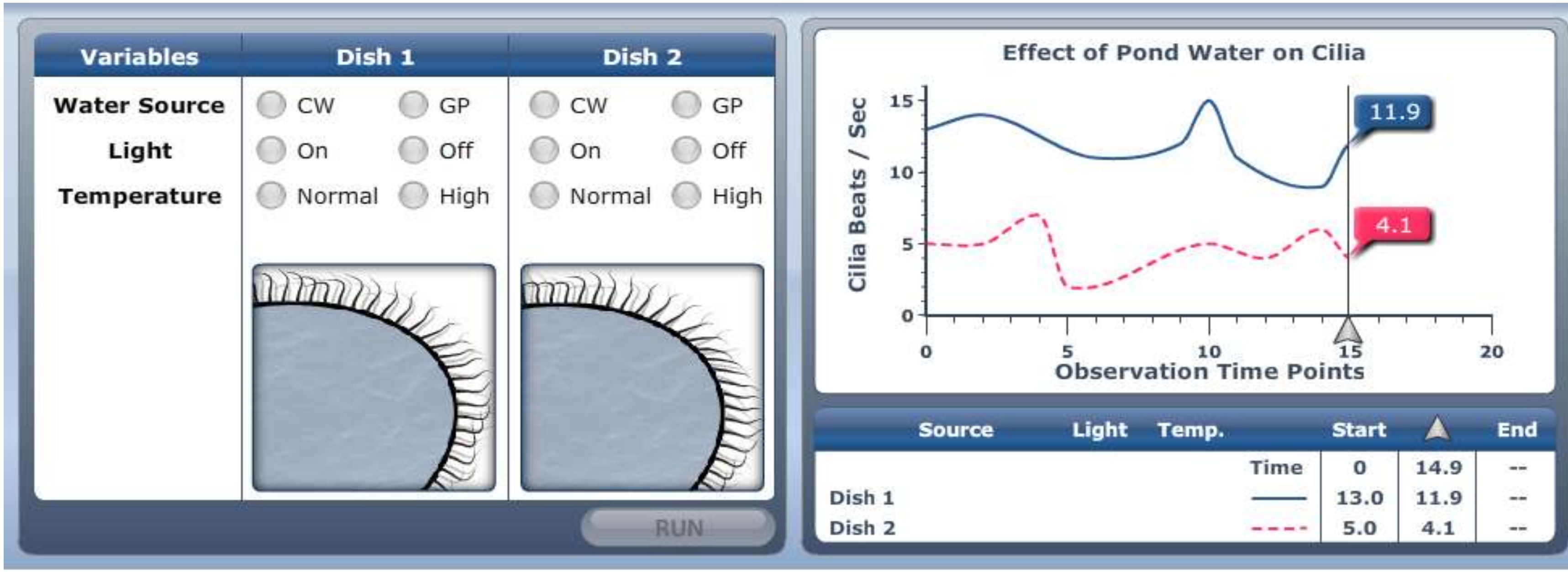
Student	Roles	Interactions	Populations	Identify	Use	Design	Conduct	Analyze	Evaluate	Communicate
Simmons85, Sara85	BB	BB	A	BB	BB	BB	B	BB	BB	A
Simmons86, Sara86	BB	BB	A	A	B	B	B	BB	BB	A
Simmons87, Sara87	BB	BB	A	A	B	B	B	BB	BB	B

SimScientists Signature Tasks

- Culminating, iconic tasks for a concept and practice
  - E.g., Draw a foodweb, use simulation of predator/prey, investigation effects of mass on force and motion
- Similar to tasks in NAEP Science, Technology and Engineering Literacy
- To permit testing of hard-to measure constructs, particularly science practices
- Several combined in a testing period

SimScientists WestEd

SimScientists
Possible Human Body Systems Signature Task: Test Effects of Pollution on Cells



You hypothesize that contaminated water in Golden Pond is slowing the movement of the Paramecia's **cilia**.

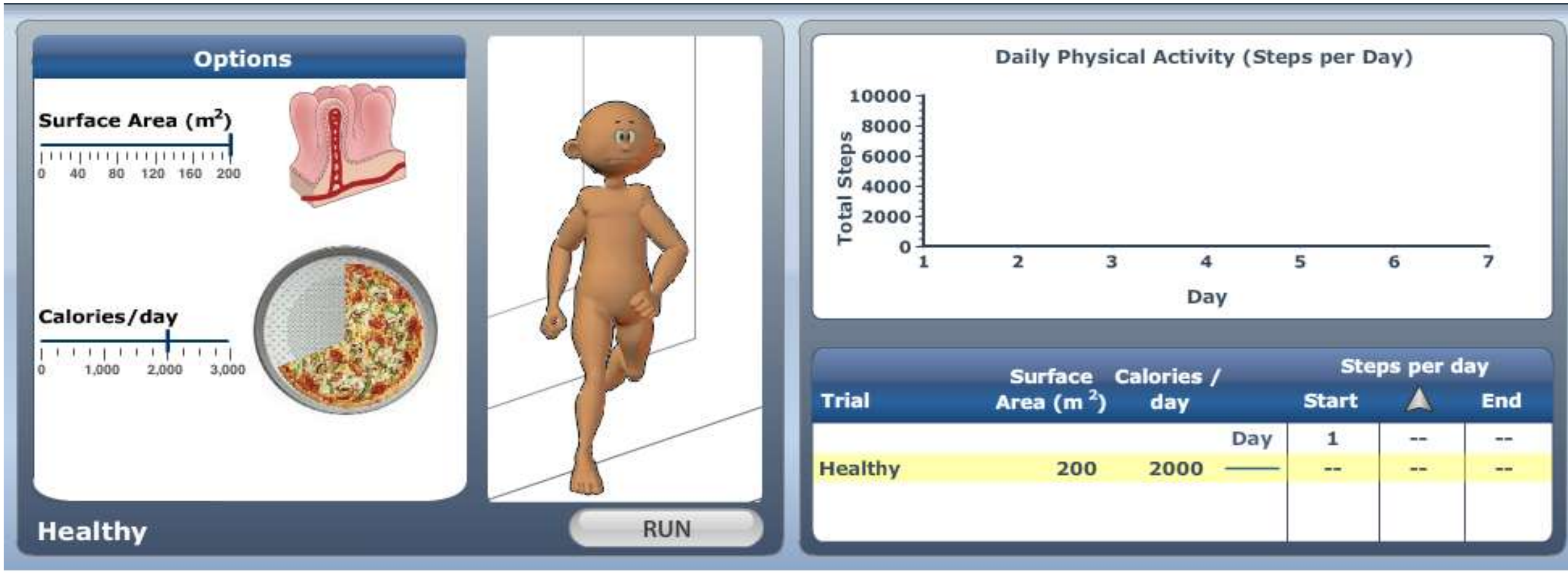
Using two dishes of Paramecia, design an experiment to test the effect of water from **Golden Pond (GP)** on cilia beats per second compared to uncontaminated water from **Clearwater Pond (CW)**.

To test your hypothesis:

- For each dish, make selections for Water Source, Light, and Temperature.
- Click **RUN** to observe the results

When you are satisfied with your experiment click **NEXT**.

SimScientists
Possible Human Body Systems Signature Task: Test Effects of Calories on Activity Level



First, observe the settings for a **healthy patient**.

The **villi surface area** is set at the maximum value (200 m²).

The **calories per day** is set at 2000. This is the total calories consumed by an average patient per day.

Click RUN to see the physical activity level of a healthy SimPatient.

SimScientists WestEd

SimScientists Research Findings

- AAAS review of alignment of content and inquiry targets with national and state standards
- Cognitive labs
- Classroom feasibility testing
- Pilot testing
- Field testing
  - EAG study involving 4 states, 28 districts, 58 teachers, 6,000 students
  - Calipers II and MASS, 3 states, 3 districts, 28 teachers, 2,500 students

SimScientists Research Findings

- Technical quality
- Implementation evaluation
- Effects of embedded on summative simulation benchmark and conventional posttest

Current Findings

The SimScientiststs simulation-based assessments

- Measure constructs not tested well by static modalities
- Can discriminate measures of inquiry and content
- The curriculum-embedded assessments seem to have positive effects on student learning
- The summative benchmark assessments have sufficient technical quality to be components of a state science assessment reporting system

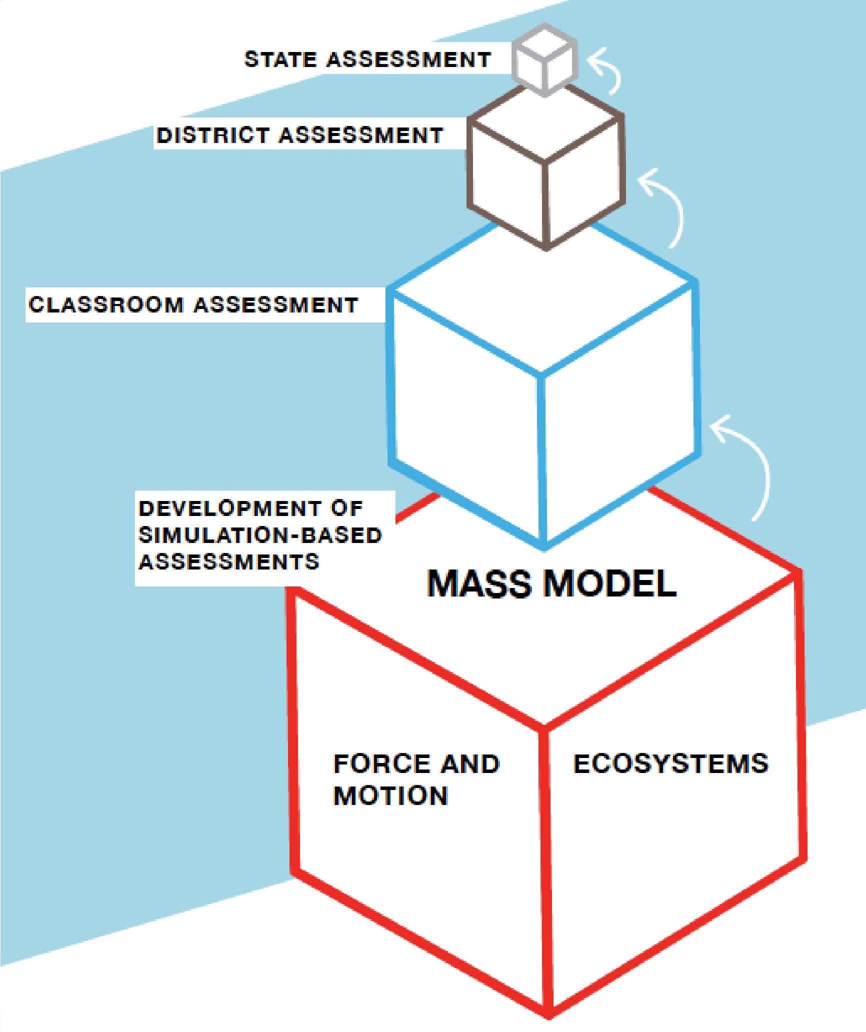
Summary of Analyses

- Correlations
  - Moderate correlations between benchmark and post test (0.57–0.64)
  - Correlations between content and inquiry are higher on the post test than the benchmark
- Gap analysis for ELLs and SWDs
  - Both groups perform better than expected on the benchmark assessment (based on their post test ability estimates)
- Reliability
  - High for all measures (coefficient alpha: 0.83–0.89)
- Effect of the treatment
  - Small, significant effect on the post test (0.07–0.08)
  - Moderate, significant effect on the benchmark (0.3–0.4)
  - Larger effect on benchmark inquiry than content (up to 0.58)

SimScientists WestEd

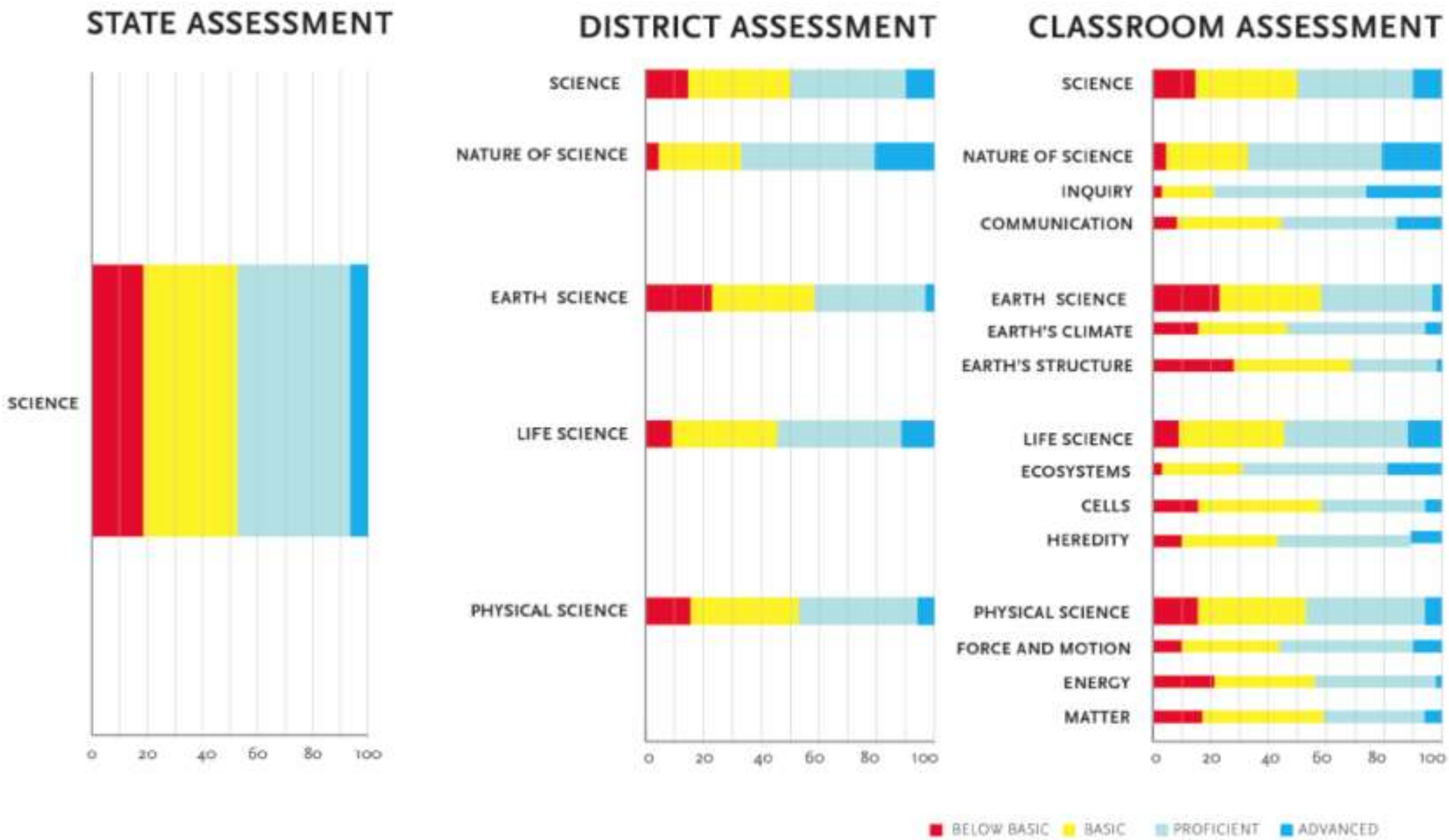
Balanced, Multilevel Assessment System Models

- Reporting benchmark results alongside district and state data
- Matrix sampling of short "signature" tasks from different topics



SimScientists WestEd

Side-by-Side Model



SimScientists WestEd

Signature Task Model

State Test Forms

Matrix Sampling

Simulation-based task item bank

Specifications and Simulation environments

Simulation-Based Classroom Assessments

SimScientists
WestEd

Continuing Research

- Study vertically aligned simulation based assessment suites for life and physical science of
  - Classroom assessments
    - curriculum embedded assessments (for formative purposes)
    - benchmark assessments (for summative purposes)
  - Large scale assessments
    - signature tasks (for summative purposes)

Summary of SimScientists Innovations

- Constructs measured-difficult or not well measured by static formats
  - System Models
  - Core concepts for a science system
  - Science Practices
- Observations
  - Observable variables from problem-based tasks and questions
  - Varied formats-selected, constructed, e.g., -explanations, arrows, set sliders, save trials
- Outcome space/interpretations
  - Proficiency calculated by Bayes Nets aggregating responses on observable variables linked to content and inquiry targets
  - Auto scoring based on rules, teacher scoring of constructed responses based following online rater training

Challenges for Scaling SimScientists Assessments

- Attention to using technologies to expand the constructs in NGSS measured, how, when, where
- Commitment to vertical integration of interactive tasks based on common specifications
- Support for development of simulation-based assessments at classroom and state level across units and grades
- Support for collections of interactive tasks
- Opportunities for students to experience simulation-based assessments as formative assessments
- Need for new measurement models for complex, interactive tasks
- Models including matrix sampling!!!

Contact information

Website:

www.simscientists.org

Email:

equellm@wested.org

SimScientists WestEd