Calipers II: Using Science Simulations to Assess Complex Science Learning

Edys Quellmalz, Barbara Buckley, Mark Loveland (WestEd) and Michael Timms (ACER)

Presented at the DRK12 Principal Investigator Meeting, Arlington, VA, June 14, 2012

Funding provided by the National Science Foundation-Grant 0733345 and the US Department of Education-OESE Grant 09-2713-126

Calipers II Goals

1. Develop simulation-based assessment modules to *supplement* and *extend* the science knowledge and skills typically addressed in static print materials
2. Foster deep learning about science systems and use of inquiry practices
3. Document the effectiveness, technical quality, feasibility, and utility of simulation-based science environments for promoting and assessing science standards

Our Theoretical Foundation

Integrates research on

- *Model-based learning* (Gobert & Buckley, 2000)
  - System Framework-components, interactions, and emergent system behavior
  - The formation, use, evaluation and revision of mental models
- *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 WestEd

Principled Assessment Design Approach (Mislevy et al., 2003)

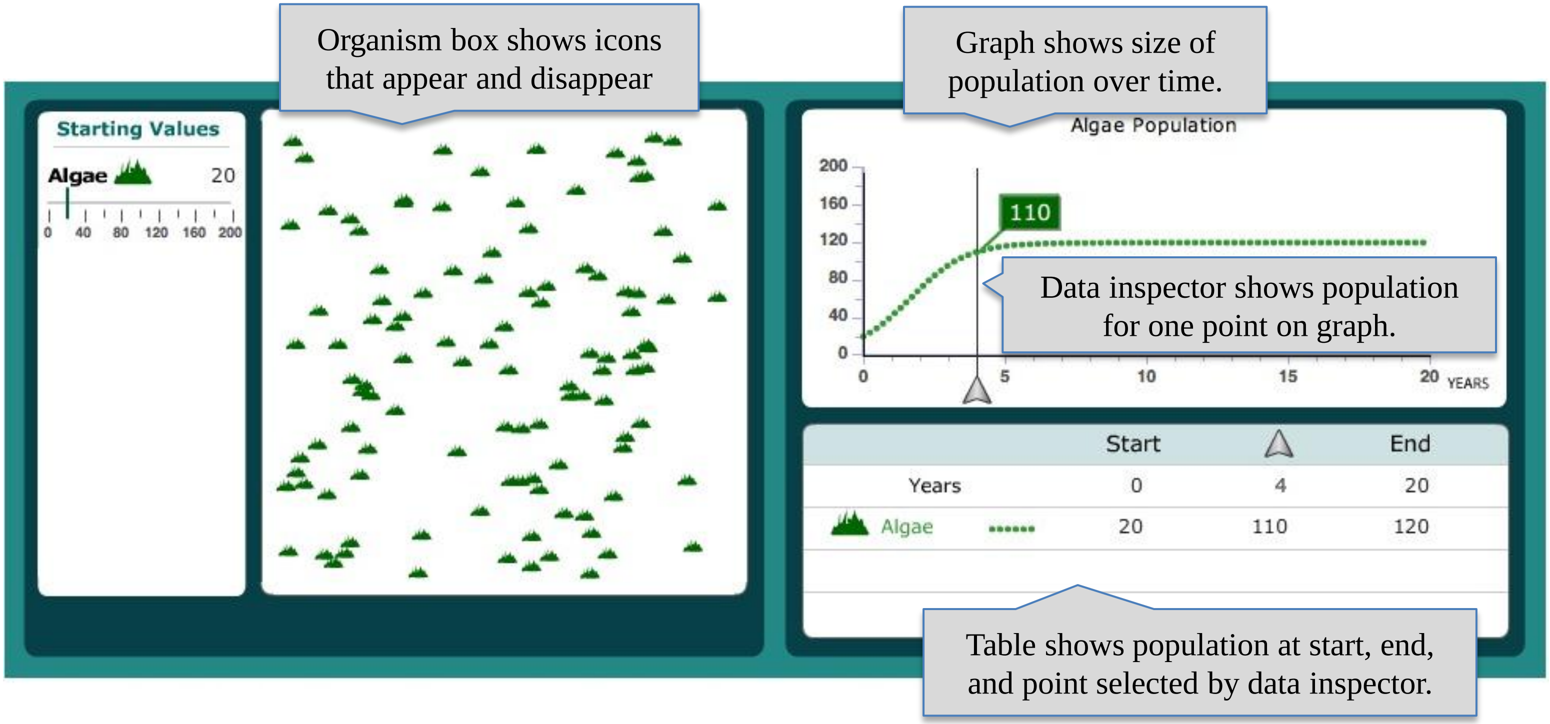
Student Model	↔	Evidence Model	↔	Task Model
What complex of knowledge, skills, or other attributes should be assessed?		What behaviors or performances should reveal the relevant knowledge and skills described in the student model?		What tasks or situations should elicit the behaviors or performances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

SimScientists WestEd

Calipers II Design Principles

- Framework of models of science systems and integrated science practices
- Authentic, problem-based inquiry
- Scaffolding
- Formative assessment
- Collaborative science practices
- Discourse for sense-making and scientific arguments

Multiple Modes of Representation
Active Inquiry



SimScientists WestEd

Calipers II Assessment Suites

1. A simulation environment representing age-appropriate models of a science system
2. Problem-driven curriculum inquiry activities
3. Embedded formative assessments with an intelligent scoring system that will provide immediate feedback
4. A coaching system tailored to individual responses and providing multiple re-teaching scaffolds

SimScientists WestEd

SimScientists Suites

5. Offline self-assessment and reflection activities to build scientific discourse and collaboration skills

6. End-of-unit benchmark assessments of progress

7. Standards-based professional development

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

What's in a Model of a Science System?

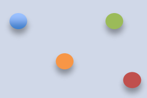
Components that have structures and rules of behavior

I**nteractions** among components and their environment, as permitted by their structures and behaviors.

Complex **behaviors or properties of the system that emerge** from these interactions.

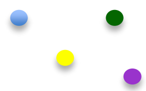
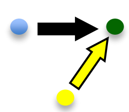
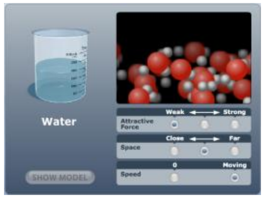
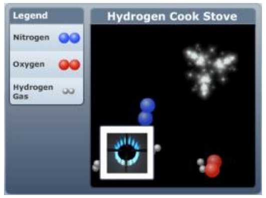
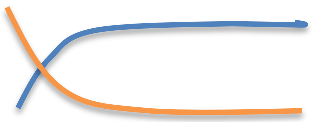
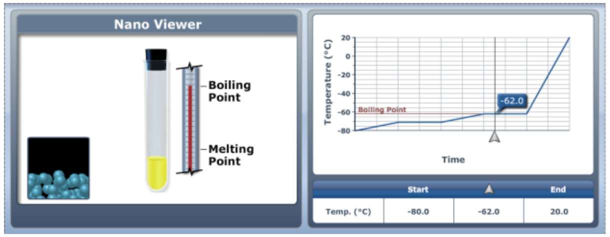
SimScientists WestEd

Ecosystem Target Model

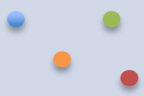
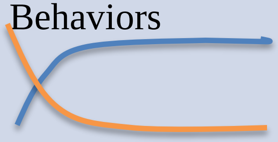
System Model Levels	Model Level Descriptions	Content Targets by Model Level	Inquiry Targets
Components 	What are the components of the system and their rules of behavior?	Every ecosystem has a similar pattern of organization with respect to the roles (producers, consumers, and decomposers) that organisms play in the movement of energy and matter through the system.	Use principles to identify role of organisms.
Interactions 	How do the the individual components interact?	Matter and energy flow through the ecosystem as individual organisms participate in feeding relationships within an ecosystem.	Observe interactions among organisms.
Emergent Behaviors 	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.	predict observe explain investigate

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

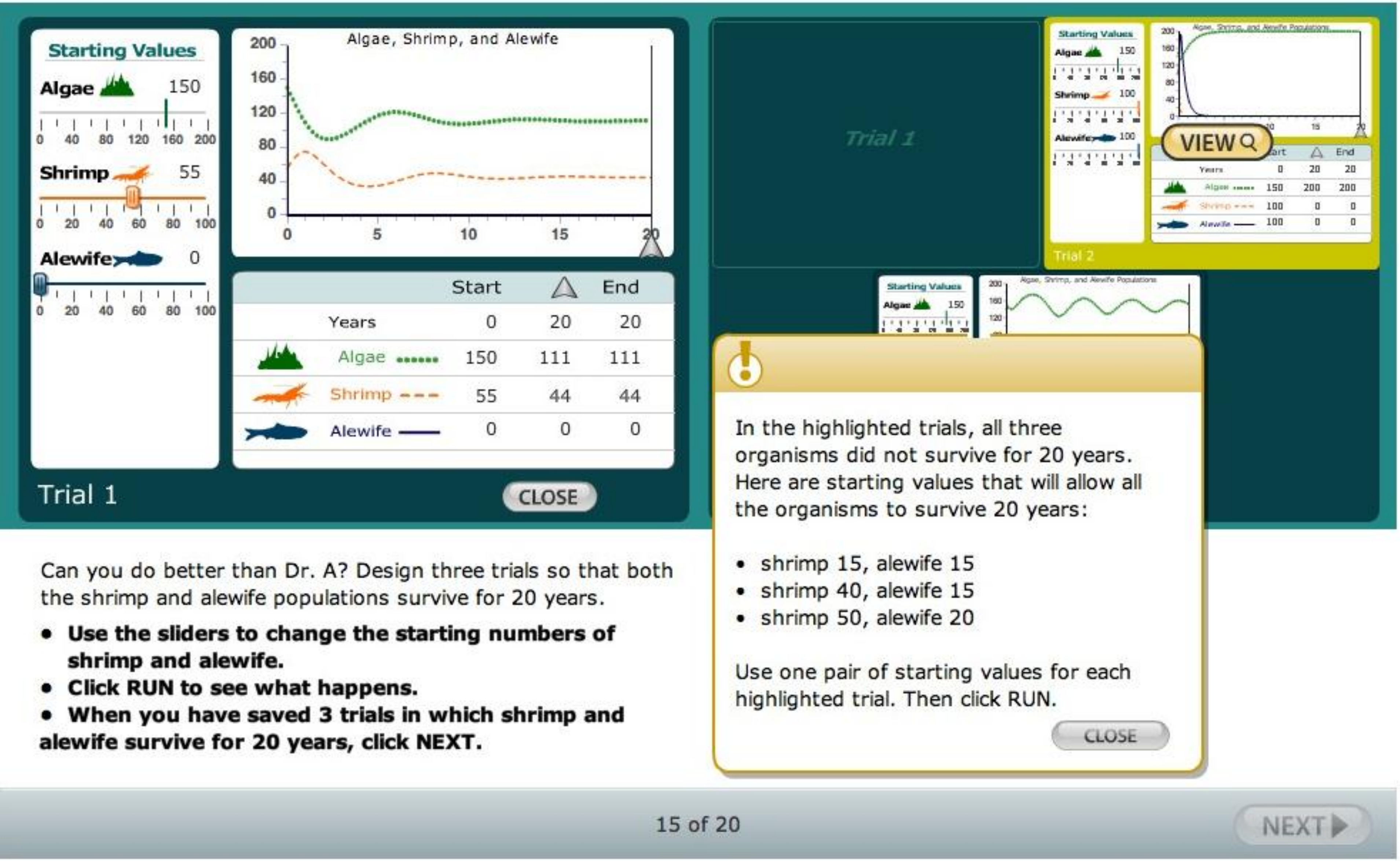
Examples System Models for Other Science Domains

Model Levels	Force & Motion	Atoms & Molecules	Climate
Components 	Objects have mass. A force is a push or pull on an object. Forces have direction and magnitude. Friction is a force that acts opposite the direction of motion.	All matter is made of particles that are in constant motion. Particles have size, shape, and structures that influence their interactions.	Earth's orbit and axial tilt strongly influence the amounts of solar energy received by different locations.
Interactions 	Forces are pushes and pulls that can affect the motion of an object. The change in motion is dependent on the strength of the forces, the direction of the forces and the mass of the object.	When particles collide, larger structures may be created. The size, shape and energy of the particles determine how the particles pack together.	This produces differential heating that results in convection currents in the atmosphere and the ocean that move energy and water around the planet.
Emergent Behaviors 	An object's motion can be described by its position, direction of motion and speed. This motion may be represented graphically.	Matter has properties that emerge from the nature, arrangement and motion of its constituent particles.	Climates in a region are strongly influenced by latitude, patterns of air and water circulation, altitude, and proximity to large bodies of water.

Demo: Embedded Assessment

Formative Assessment Features

- Immediate, individualized feedback and coaching
- Progress reports for students and teachers
- Reflection activities that address students' needs, promote transfer, and scientific discourse
- Timely information that teachers can use to tailor and adjust instruction


Starting Values
Algae 150
Shrimp 55
Alewife 0
Algae, Shrimp, and Alewife
Start
End
Years 0 20 20
Algae 150 111 111
Shrimp 55 44 44
Alewife 0 0 0
Trial 1
CLOSE
Trial 1
VIEW
Trial 2
In the highlighted trials, all three organisms did not survive for 20 years. Here are starting values that will allow all the organisms to survive 20 years:
• shrimp 15, alewife 15
• shrimp 40, alewife 15
• shrimp 50, alewife 20
Use one pair of starting values for each highlighted trial. Then click RUN.
CLOSE
Can you do better than Dr. A? Design three trials so that both the shrimp and alewife populations survive for 20 years.
• Use the sliders to change the starting numbers of shrimp and alewife.
• Click RUN to see what happens.
• When you have saved 3 trials in which shrimp and alewife survive for 20 years, click NEXT.
15 of 20
NEXT

Progress Reports to Students (Formative)

Back to Home

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

Progress Reports to Teachers (Formative)

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	NH OT	12 (46%)	4 (15%)	10 (38%)
Interactions	NH OT	15 (58%)	4 (15%)	7 (27%)
Inquiry				
Identifying	NH OT	15 (58%)	5 (19%)	6 (23%)
Using	NH OT	10 (38%)	5 (19%)	11 (42%)

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

SimScientists WestEd

Grouping Recommendations for Classroom Reflection Activity (Formative)

ASSESSMENT: Mountain Lake - Food Web

CLASS: Period 7

Go!

NH Needs Help P Making Progress OT On Track

Reflection Activity PDF

Group A students needed little help on either roles or interactions
Group B students needed help with interactions, but not with roles.
Group C students needed help with understanding the roles of organisms in an ecosystem.

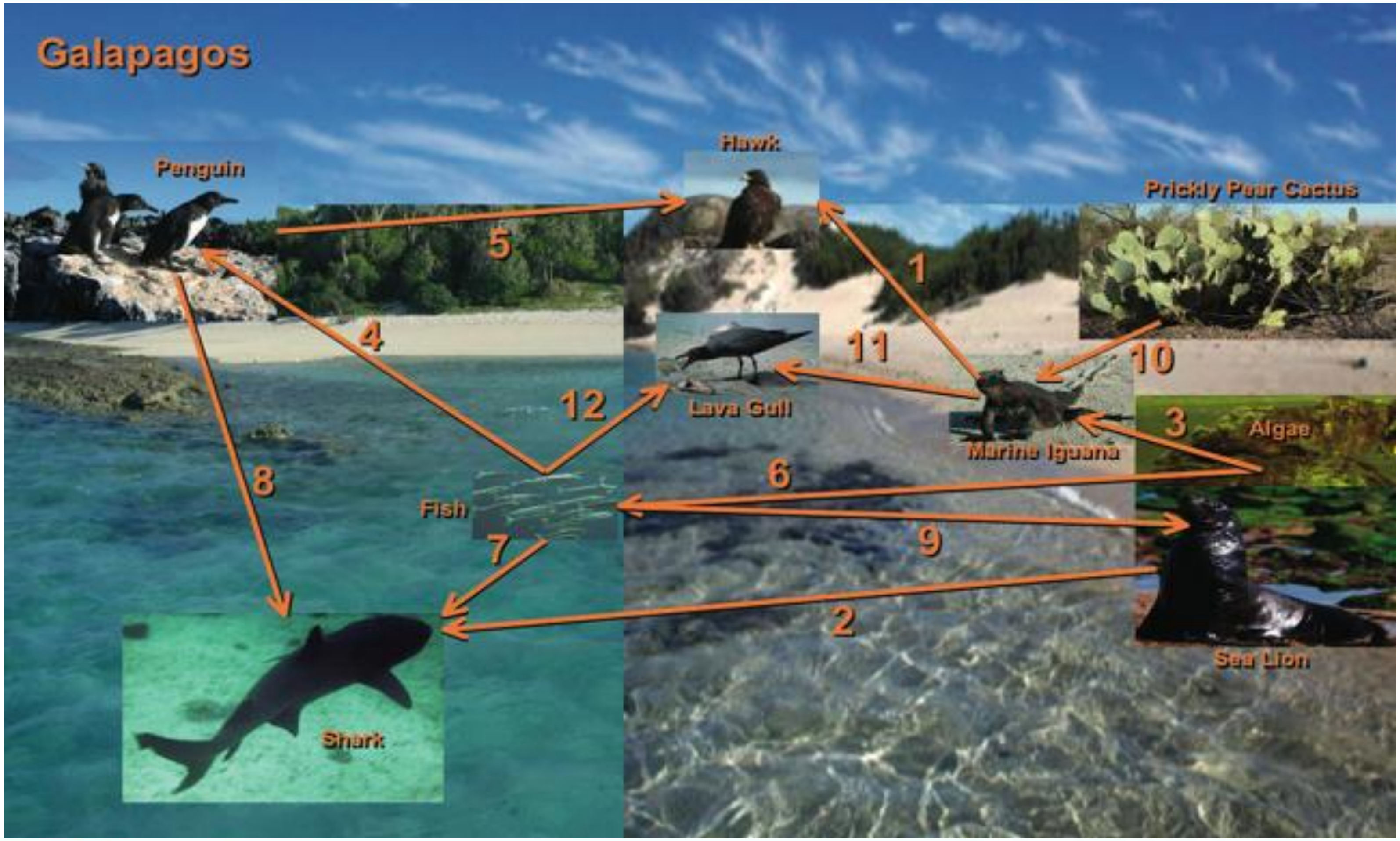
Student	Refl Gr.	Roles	Interactions	Identifying	Using
Student 1	C	P	NH	NH	OT
Student 1	C	NH	NH	NH	NH
Student 3	A	OT	OT	OT	OT
Student 4	A	OT	OT	OT	OT
Student 5	C	NH	NH	NH	NH
Student 6	C	NH	NH	NH	P
Student 7	C	P	NH	NH	P
Student 8	C	NH	NH	NH	NH
Student 9	C	NH	OT	NH	P
Student 10	B	OT	NH	OT	P

SimScientists WestEd

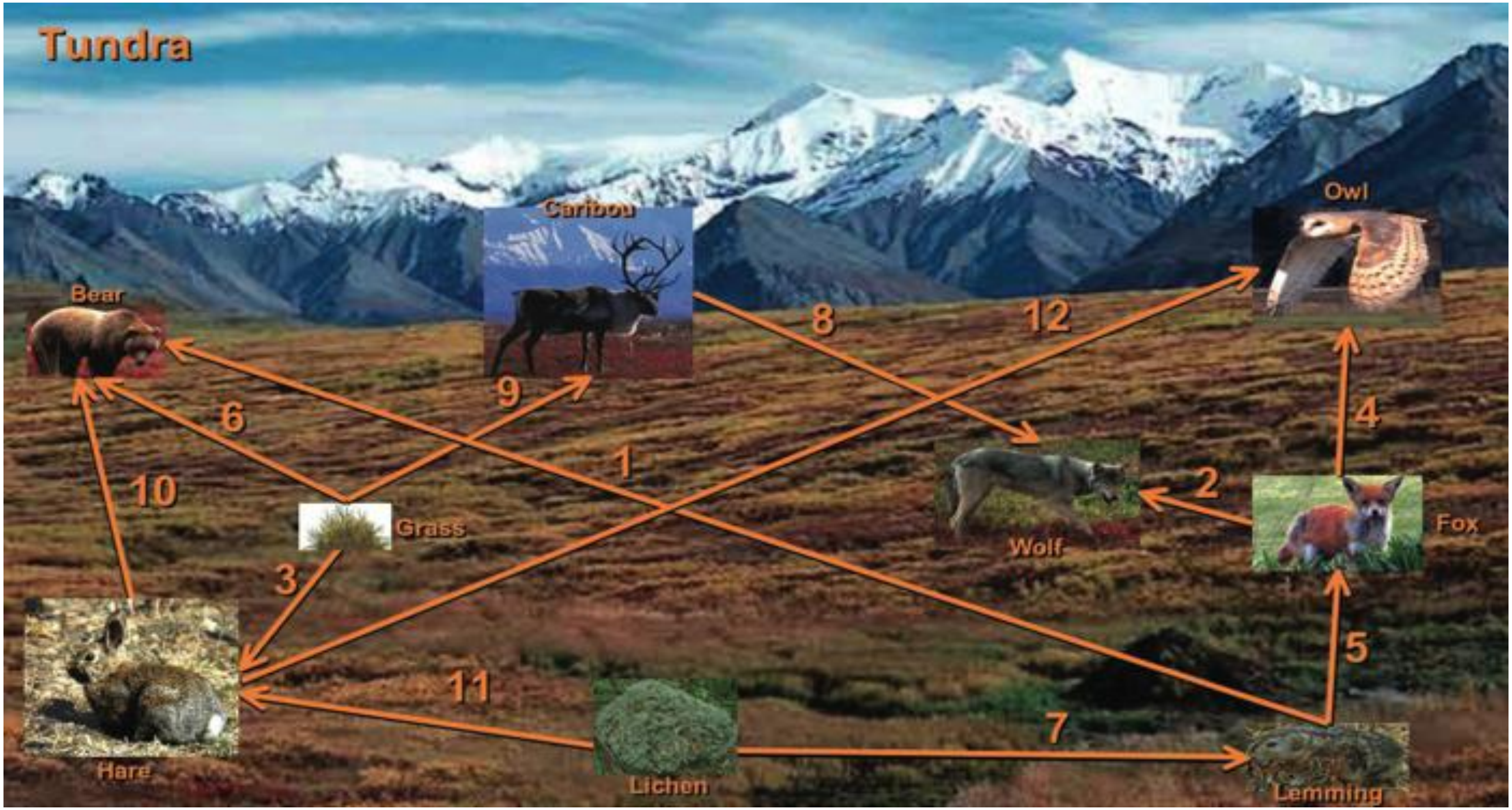
Classroom Reflection Activity (Adjusting Instruction)

- Formative use of assessment results
  - Students assigned to teams based on embedded results
- Transfer to different, more complex system
- Jigsaw structure
  - Allows differentiated instruction via tasks of varying difficulty
  - Promotes small and large group discourse and collaboration
- Guidance for teacher
  - Teacher review of key points in simulation
  - What to look for during group work and questions to pose in response
  - Posters and presentations
  - Evaluation of posters and presentations by students and teachers

Transfer to New, More Complex Ecosystem



Calipers II
Reflection Activities: Ecosystems



SimScientists WestEd

Ecosystem Benchmark Assessment: (Summative) Assess Transfer to New Ecosystem

SimScientists WestEd

Summary Benchmark Report (Summative)

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%)
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.					
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%)

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 WestEd

Calipers II Data and Findings

- AAAS review of alignment of content and inquiry targets with national and state standards
- Cognitive labs
- Classroom feasibility testing
- Pilot testing
- Field testing
  - 4 states, 28 districts, 58 teachers, 6,000 students
- Findings
  - Technical quality
  - Implementation evaluation
  - Effects of embedded on summative simulation benchmark and conventional posttest

EAG State Design Panel Partners

Pilot states

- Nevada (lead state)
- North Carolina
- Utah

Advisory states

- Massachusetts
- Connecticut
- Vermont

CRESST Evaluation Conclusions

- Teachers were able to implement as intended.
- Observations found that students were actively engaged.
- Teachers and students generally believed that the SimScientists program was beneficial to learning.
- Teachers found the automatically scored, immediate feedback—especially the reports generated by the questions—helpful to students. The instant reports allowed teachers to easily see which questions students had the most difficulty with so that they could tailor their lessons accordingly.

CRESST Evaluation Conclusions

- Teachers collectively agreed that the simulation-based assessments had greater benefits than traditional paper-and-pencil tests because of the instant feedback, interaction, and visuals.
- Teachers agreed that the assessments would be useful in measuring their individual state standards.

Analyses of Performance

- Compare performance on simulation-based assessments to traditional assessments
- Compare performance for all students, for English Learners, and for students with disabilities

Correlations Between Posttest and Benchmark Ability Estimates

Topic	Content	Inquiry
Ecosystems	0.64	0.57
Force and Motion	0.61	0.60

SimScientists WestEd

Correlations Between Content and Inquiry Ability Estimates for Each Assessment

Topic	Posttest	Benchmark
Ecosystems	0.85	0.70
Force and Motion	0.92	0.80

SimScientists WestEd

Gaps in Total Performance Between English Learners and the General Population

Topic	Posttest	Benchmark
Ecosystems	24.0% (n=123)	10.6% (n=126)
Force and Motion	27.4% (n=50)	13.6% (n=50)

SimScientists WestEd

Gaps in Inquiry Skills Performance Between English Learners and the General Population

Topic	Posttest	Benchmark
Ecosystems	25.6% (n=123)	6.6% (n=126)
Force and Motion	35.1% (n=50)	10.9% (n=50)

Gaps in Total Performance Between students with Disabilities and the General Population

Topic	Posttest	Benchmark
Ecosystems	25.5% (n=183)	5.6% (n=189)
Force and Motion	20.3% (n=153)	6.2% (n=153)

SimScientists WestEd

Gaps in Inquiry Skills Performance Between Students with Disabilities and the General Population

Topic	Posttest	Benchmark
Ecosystems	20.2% (n=183)	8.4% (n=189)
Force and Motion	15.7% (n=153)	7.0% (n=153)

SimScientists WestEd

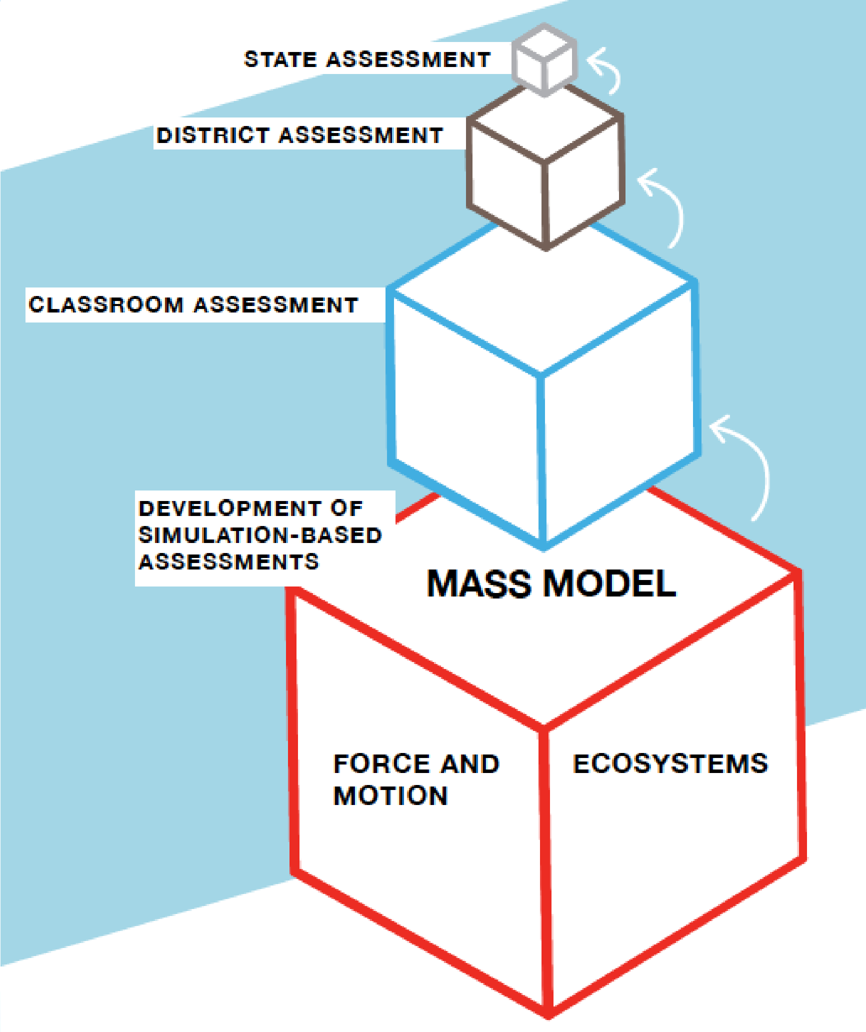
Calipers II Field Testing (in progress)

- To examine effects of the curriculum-embedded assessments on learning
- Randomized control and treatment groups
  - With and without curriculum-embedded assessments
  - Partial results-Ecosystems
    - 763 students, 5 teachers (full sample 21 T, ~2400 students)
    - Effect size on pre-post conventional AAAS items 0.19
    - Effect size simulation based benchmark 0.43
  - Full field test in progress for Ecosystems and Atoms and Molecules

SimScientists WestEd

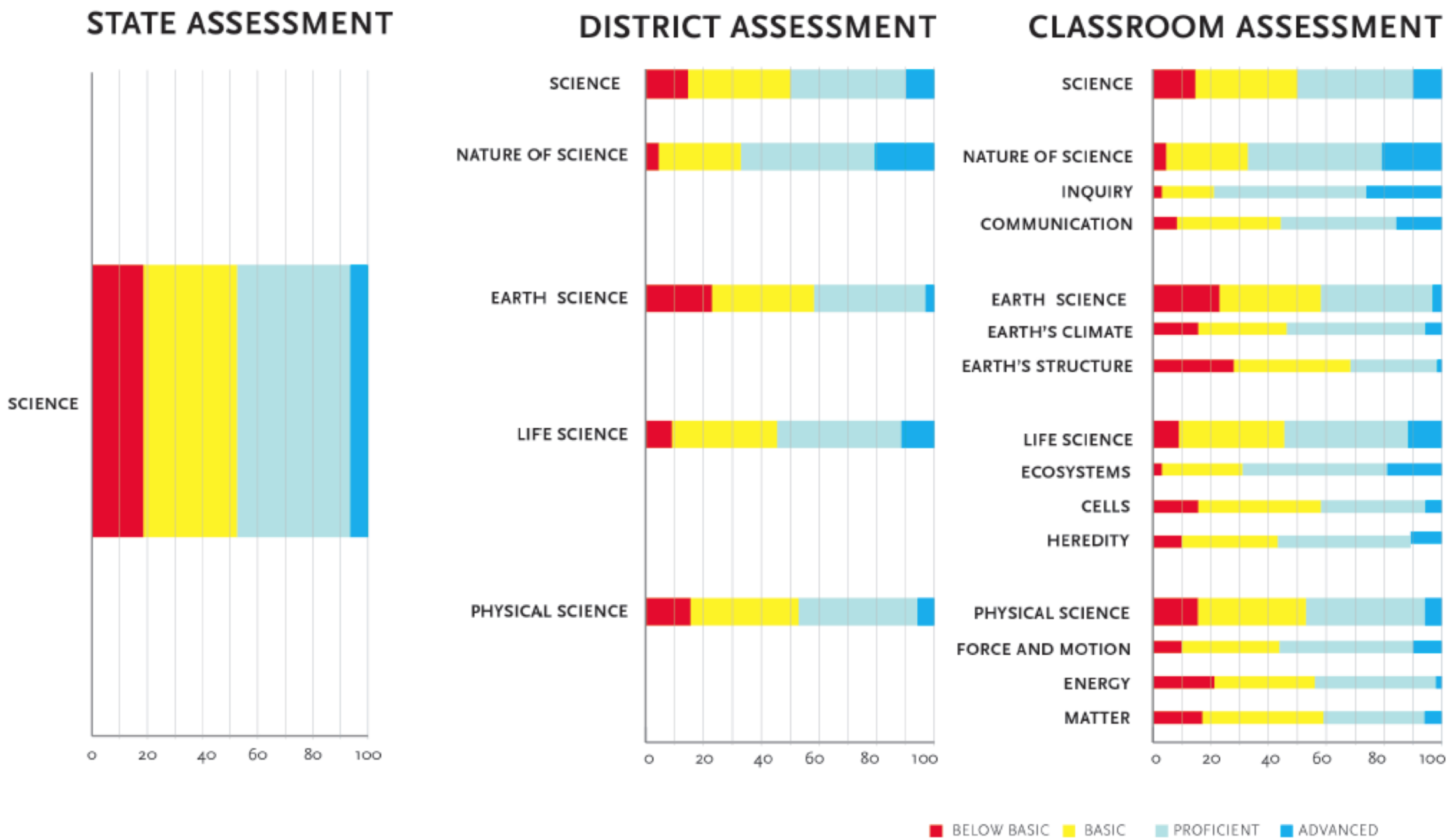
Balanced, Multilevel Assessment System Models

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



SimScientists WestEd

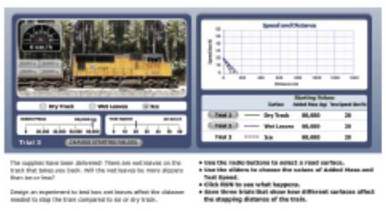
Side-by-Side Model

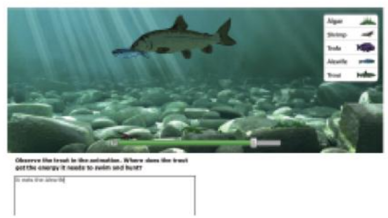


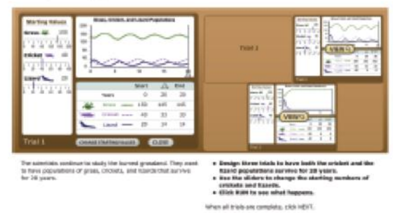
Signature Task Model

State Test Forms

Matrix Sampling

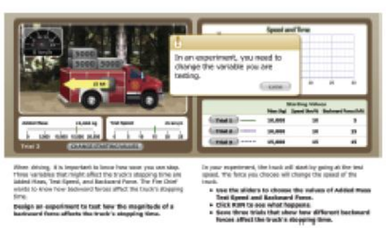


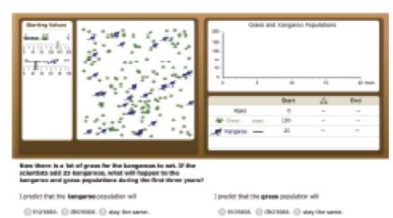


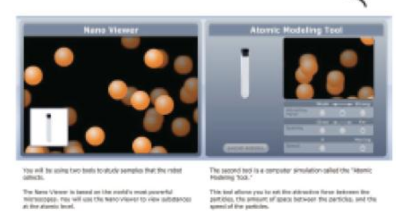


Simulation-based task item bank

Specifications and Simulation environments







Simulation-Based Classroom Assessments

SimScientists WestEd

Current Findings

The Calipers II simulation-based assessments

- Measure constructs not tested well by static modalities
- Can provide separate 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

Challenges

- Scheduling computer access
- Convincing psychometricans
- Progress from prototypes to developing grade-band suites
- Replications of documentation of technical quality
- Delivery across multiple technical infrastructures
- Models for scaling
- Models for sustainability

Next Steps

- Study vertically aligned simulation based assessment suites of
  - Classroom assessments
    - curriculum embedded assessments (for formative purposes)
    - benchmark assessments (for summative purposes)
  - Accountability assessments
    - signature tasks (for summative purposes)

SimScientists WestEd

Potential Collaborations

- Use of developed Calipers assessments as one of your external outcome measures
- Development of assessments embedded within your programs
- Co-development of new simulation-based learning and assessments

Contact Information

Website:

www.simscientists.org

Email:

equellm@wested.org

SimScientists WestEd