SimScientists: Affordances of Science Simulations for Formative and Summative Assessment

Edys Quellmalz and Matt Silberglitt
WestEd

Presented at the Maryland Assessment Conference, October 16, 2014

This material is based upon work supported by grants awarded to WestEd from the National Science Foundation (DRL-0733345, -1020264, -1221614, -0814776), and the US Department of Education (R305A080225, R305A120390, R305A080614, R305A130160), and a grant awarded to the Nevada Department of Education from the US Department of Education (09-2713-126). Any opinions, findings, 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.

COPYRIGHT WESTED, 2013

SimScientists Team

Edys Quellmalz, PI
Daniel Brenner
Barbara Buckley
Jodi Davenport
Mark Loveland
Matt Silberglitt
Arthur Sussman
Mike Timms, ACER

Teacher co-developers

Kevin Jordan
Kevin Huang
Michelle Lamar
Ellen Mandinach

Andrew Grillo-Hill
Marina Varfolomeeva
Jackie Powers
May Miller Ricci
Kim Luttgen

SimScientists WestEd

COPYRIGHT WESTED, 2013

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 WestEd
COPYRIGHT WESTED, 2013

SimScientists Assessment Suites

(Curriculum-embedded Assessments and Unit Benchmark)

Life Science

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

Physical Science

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

Earth Science

- Plate tectonics
- Climate

SimScientists WestEd

COPYRIGHT WESTED, 2013

Research Foci

- ***Assessment validity and quality***
  - Alignment with the NGSS an, 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
  - Valuable 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

COPYRIGHT WESTED, 2013

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

COPYRIGHT WESTED, 2013

Next Generation Science Standards Addressed in Life Science Examples

Crosscutting concepts

- System and system models
- Energy and matter: Flows, cycles, and conservation
- 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 and carrying out investigations
- Analyzing and interpreting data
- Constructing explanations
- Engaging in argument from evidence

COPYRIGHT WESTED, 2013

Ecosystems Target Model

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.

Model Progressions

- Simulation-based assessments address progressive understanding across topics of
  - *Crosscutting concepts*
    - *E.g., Systems and System Models*

      System Framework: components, interactions, and emergent system behavior

COPYRIGHT WESTED, 2013

Learning Trajectories

- Simulation-based assessments address progressive understanding of core ideas for a particular system applied to real life problems using the science and engineering practices
  - *Core Ideas:*
    - *E.g., Human Body Systems, Ecosystems*
  - *Science and Engineering Practices:*
    - *E.g., Building and using models, planning, conducting, analyzing investigations, constructing explanation, critiquing arguments*

COPYRIGHT WESTED, 2013

Design Principles from Cognitive Science and Multimedia Learning Literature

- *Cognitive science*
  - Types of knowledge
    - declarative, procedural, schematic, strategic
    - integration of knowledge
    - active application of processes
- *Multimedia research*
  - *E.g., minimize extraneous processing*
  - *Attention guiding, cueing, signaling*
  - *Coherence, contiguity*
  - *Progressive complexity, segmentation*
  - *User control*
  - *Modality*

COPYRIGHT WESTED, 2013

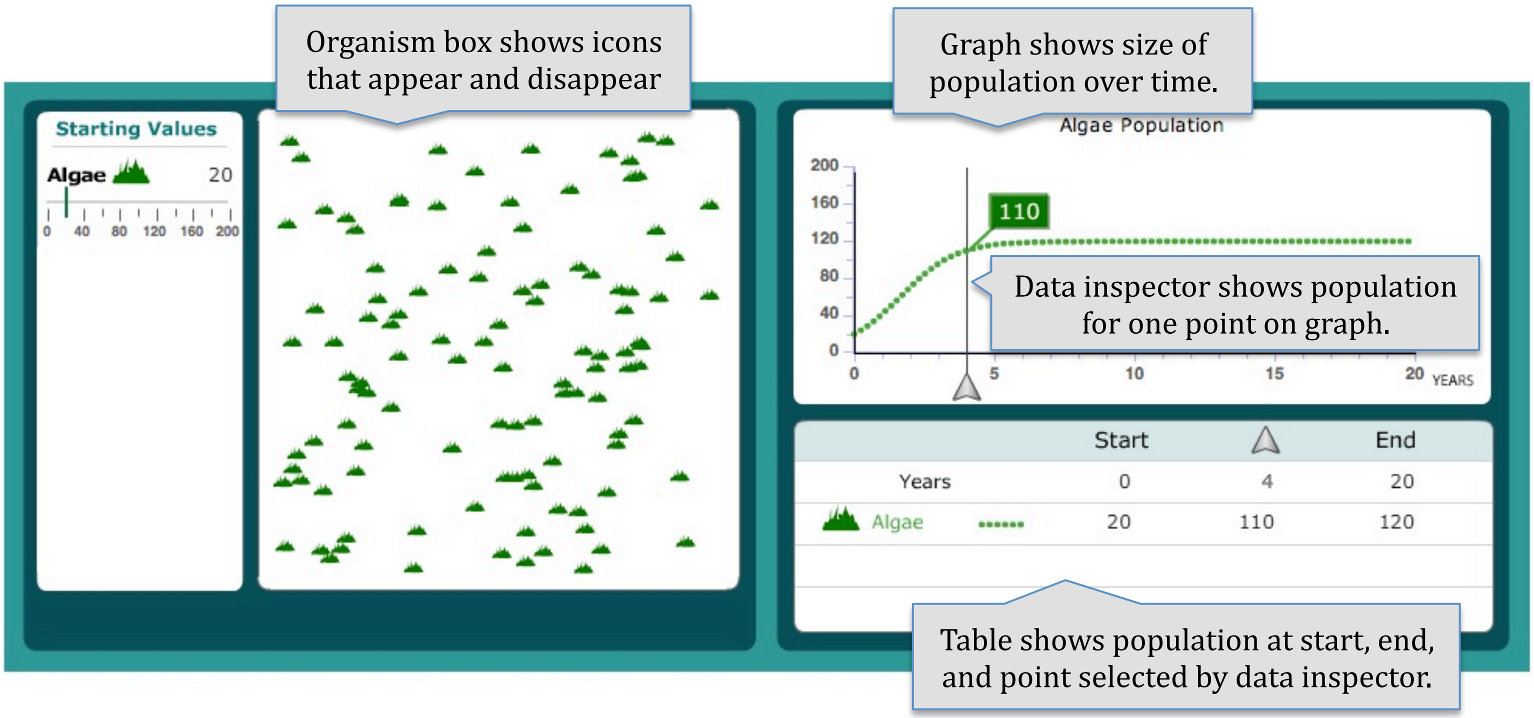
Technology Affordances

- Animations of dynamic science system phenomena
  - Can observe, review, gather evidence
- Simulation-based investigations
  - Permit multiple trials
  - Permit saving trials
  - Data inspector
  - Explicit responses and log file data (e.g., sequence, time)
- Multiple, simultaneous representations
- Highlighting
- Zoom
- Drawing

SimScientists WestEd

COPYRIGHT WESTED, 2013

Figure 2. Multiple modes of representation



COPYRIGHT WESTED, 2013

Task Models

Simulation Environments — for science systems

physical science, life science, earth science

Benchmark Assessments — one per topic

summative

end-of-unit

one period

Embedded Assessments —two to three per topic

formative

inserted during unit

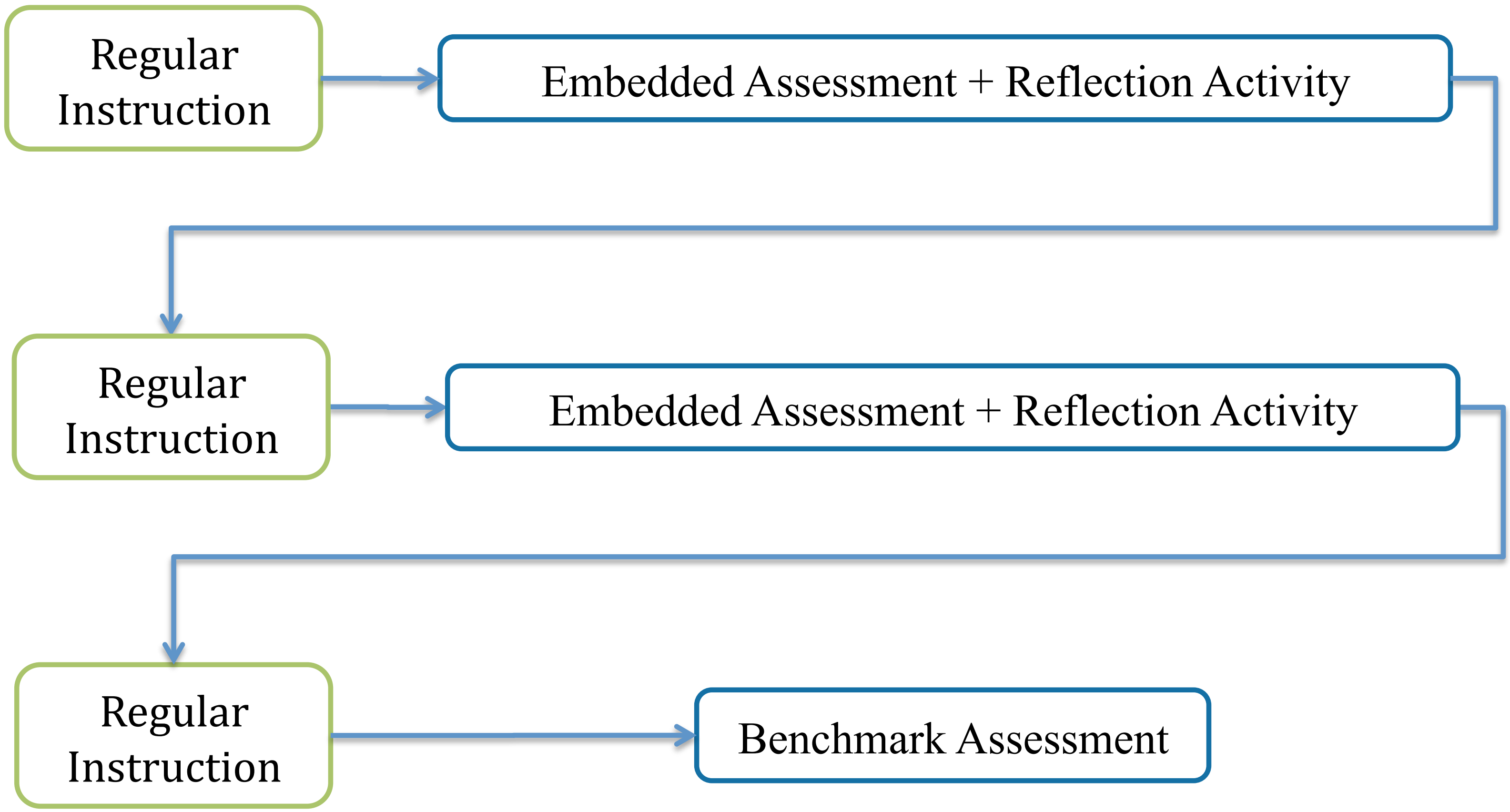
one period

SimScientists WestEd

COPYRIGHT WESTED, 2013

SimScientists Assessments

Embedded & Benchmark



SimScientists WestEd

COPYRIGHT WESTED, 2013

Task Models

Embedded Assessments

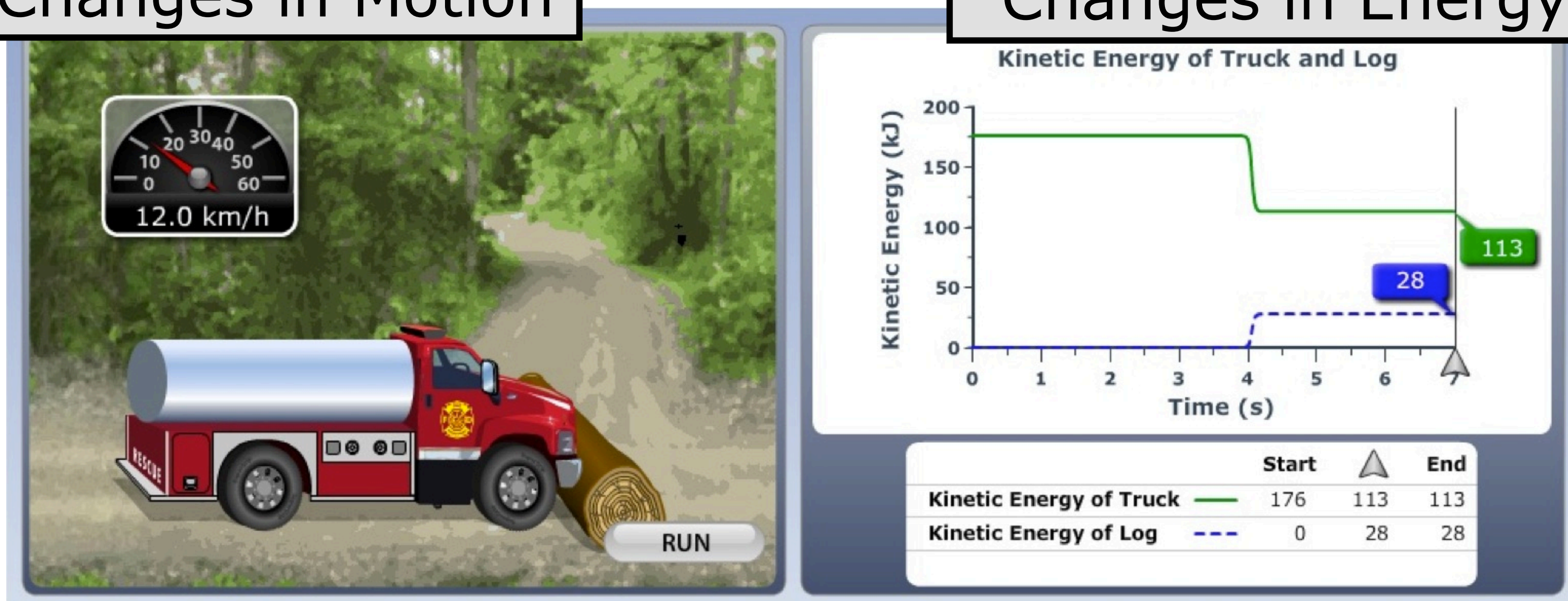
- Immediate feedback
  - Autoscored
  - Self-assessment
    - criteria
    - sample appropriate response
- Customized, graduated coaching
- Off-line self-assessment/reflection activities for adjusting instruction, transfer, collaboration, communication

COPYRIGHT WESTED, 2013

Task Model:
Example from Energy Embedded Assessment

Changes in Motion

Changes in Energy



Energy Transfer?

COPYRIGHT WESTED, 2013

Task Types

- Explain
- Describe evidence
- Justify or critique claims
- Revise and evaluate (embedded only)
- Construct/use models
- Analyze data
- Predict/observe/explain
- Predict/test/evaluate
- Design investigations
  - Design/revise/evaluate
  - Conduct/analyze/conclude
- Design solutions

SimScientists WestEd

COPYRIGHT WESTED, 2013

Response Types

- Multiple choice
- Multiple yes/no

- Constructed response
  - Text
  - Draw arrows
  - Drag and drop
  - Manipulate variables
  - Generate graphical displays
  - Produce multiple solutions
  - Save solutions

SimScientists WestEd
COPYRIGHT WESTED, 2013

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

COPYRIGHT WESTED, 2013

SimScientists Assessments

Curriculum-Embedded Assessment
DEMO

COPYRIGHT WESTED, 2013

Formative Assessment Features within the Simulation-based Embedded Assessments

- Immediate feedback
  - On core ideas/misconceptions
  - On practices
- Graduated coaching
- Progress Report
  - On Track, Making Progress, Needs Help by content and practices targets
- For individual student
- Teachers
  - Class summary with drill-down into student detail
  - Used to suggest teams and groups for jig-saw structured reflection activities that adjust instruction and support collaboration and discourse

SimScientists WestEd

COPYRIGHT WESTED, 2013

Sample Progress Reports to Students Ecosystems

SimScientists WestEd

Hi Sara | Log Out | Your Account

Report for Grasslands - Food Web

life science

Completed on 03/22/2011

Roles PROGRESSING	Every ecosystem has a similar pattern of organization with respect to the roles that organisms play in the movement of energy and matter through the system. Every ecosystem has organisms that play the roles of Producers, Consumers, and Decomposers.
Interactions PROGRESSING	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.
Identifying 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."
Using ON TRACK	Using Science Principles focuses on the ability to use patterns in observations and theoretical models to predict and explain observations that they make now or that they will make in the future. The practices assessed in this category draw primarily on schematic knowledge or "knowing why."

Copyright © 2010. All rights reserved.

Privacy Policy | Disclaimer

SimScientists WestEd

COPYRIGHT WESTED, 2013

Embedded Progress Report: Class Summary Ecosystems

Summary Report: Grasslands - Food Web

Try it | Detailed Report

ASSESSMENT: Grasslands - Food Web

CLASS: Period 4

Go!

Content		NH Needs Help	P Making Progress	OT On Track
▸ Roles	NH PROG OT	4 (17%)	6 (25%)	14 (58%)
▸ Interactions	NH PROG OT	15 (63%)	5 (21%)	4 (17%)
Inquiry				
▸ Identifying	NH OT	9 (38%)	3 (13%)	12 (50%)
▸ Using	NH PROG OT	8 (33%)	5 (21%)	11 (46%)

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

SimScientists WestEd

COPYRIGHT WESTED, 2013

Embedded Progress Report for Ecosystems: Student Details

Detailed Report: Grasslands - Food Web

Try it | Summary Report

ASSESSMENT: Grasslands - Food Web

CLASS: Period 4

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	NH	NH	NH	P
Student 2	B	OT	NH	OT	NH
Student 3	C	NH	NH	NH	NH
Student 4	C	P	NH	NH	P
Student 5	C	P	NH	P	P
Student 6	B	OT	NH	P	P
Student 7	B	OT	P	OT	OT
Student 8	A	OT	OT	OT	OT
Student 9	A	OT	OT	OT	OT

SimScientists WestEd

COPYRIGHT WESTED, 2013

Classroom Reflection Activity

- 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

COPYRIGHT WESTED, 2013

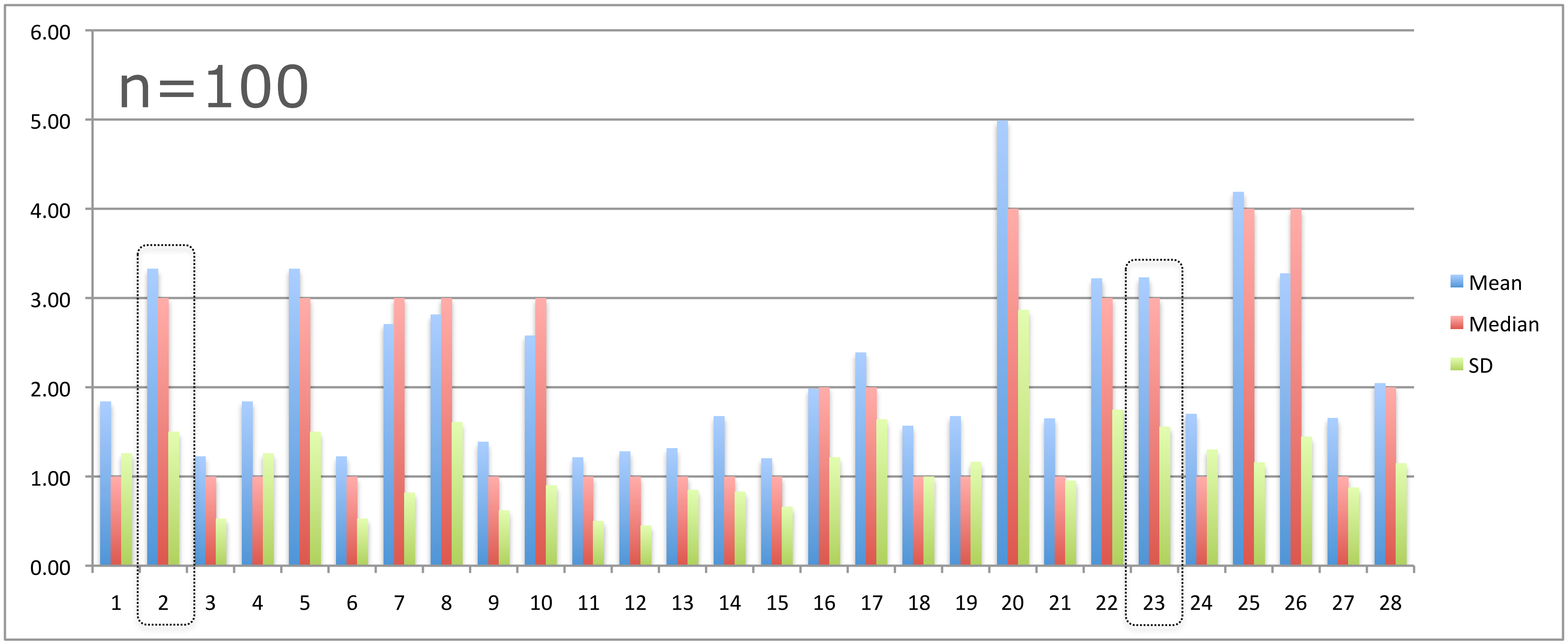
Analyses

- Descriptive statistics
- Classical psychometrics
- IRT
- Bayes' Nets
- Exploratory data mining

SimScientists WestEd

COPYRIGHT WESTED, 2013

Embedded Assessment 1
Number of Tries to Correct



COPYRIGHT WESTED, 2013

Embedded Assessment 1 Proportion Correct on Try 1

Item 1

------

item:1 (Motion.force)
Cases for this item 98 Discrimination 0.07
Item Threshold(s): 1.37 Weighted MNSQ 1.04
Item Delta(s): 1.37

Label	Score	Count	% of tot	Pt Bis	t (p)	PV1Avg:1	PV1 SD:1
0	0.00	57	58.16	-0.07	-0.72(.471)	1.01	0.46
1	1.00	41	41.84	0.07	0.72(.471)	1.10	0.62

Item 2

------

item:2 (Motion.engine)
Cases for this item 98 Discrimination 0.24
Item Threshold(s): -2.28 Weighted MNSQ 0.97
Item Delta(s): -2.28

Label	Score	Count	% of tot	Pt Bis	t (p)	PV1Avg:1	PV1 SD:1
0	0.00	4	4.08	-0.24	-2.39(.019)	0.43	0.73
1	1.00	94	95.92	0.24	2.39(.019)	1.07	0.51

SimScientists WestEd

COPYRIGHT WESTED, 2013

Embedded Assessment 1
Point Biserial for Try 1

```
Item 1
------
item:1 (Motion.force)
Cases for this item      98   Discrimination  0.07
Item Threshold(s):     1.37    Weighted MNSQ   1.04
Item Delta(s):       1.37
------------------------------------------------------------------------------
 Label     Score      Count    % of tot   Pt Bis      t  (p)    PV1Avg:1 PV1 SD:1
------------------------------------------------------------------------------
   0       0.00        57       58.16    -0.07    -0.72(.471)  1.01      0.46
   1       1.00        41       41.84     0.07     0.72(.471)  1.10      0.62
==============================================================================

Item 2
------
item:2 (Motion.engine)
Cases for this item      98   Discrimination  0.24
Item Threshold(s):    -2.28    Weighted MNSQ   0.97
Item Delta(s):      -2.28
------------------------------------------------------------------------------
 Label     Score      Count    % of tot   Pt Bis      t  (p)    PV1Avg:1 PV1 SD:1
------------------------------------------------------------------------------
   0       0.00         4        4.08    -0.24    -2.39(.019)  0.43      0.73
   1       1.00        94       95.92     0.24     2.39(.019)  1.07      0.51
==============================================================================
```

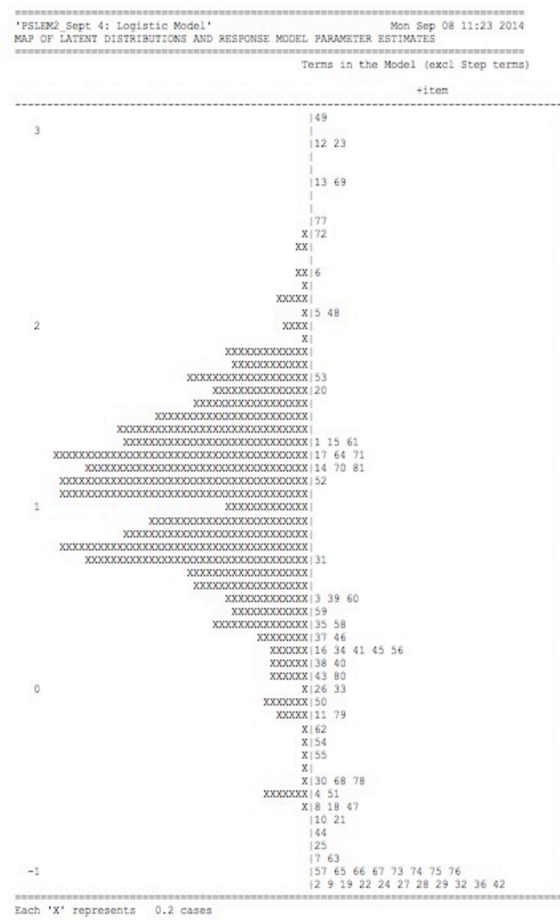
SimScientists WestEd

COPYRIGHT WESTED, 2013

Embedded Assessment 1 IRT Analysis and Wright Map

TERM 1: item

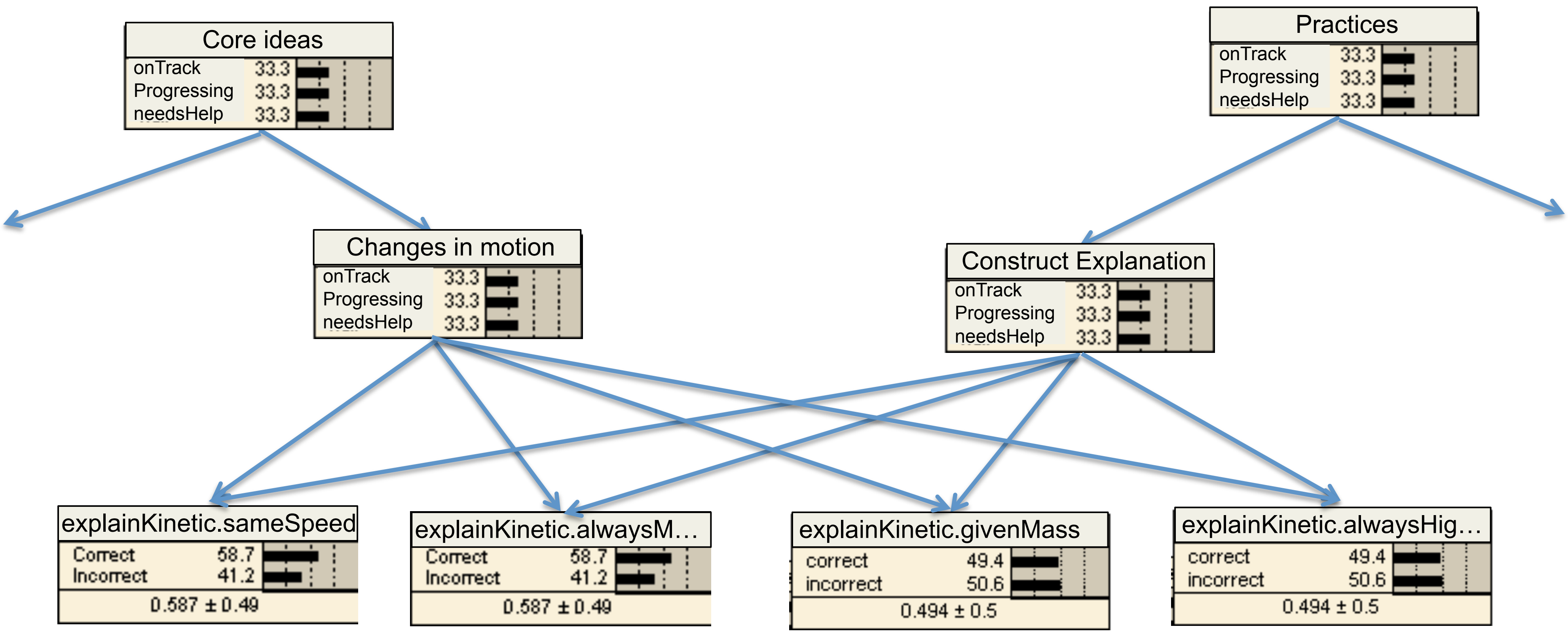
VARIABLES				UNWEIGHTED FIT			WEIGHTED FIT		
	item	ESTIMATE	ERROR	MNSQ	CI	T	MNSQ	CI	T
1	Motion.forc..	1.366	0.217	1.09	(0.72, 1.28)	0.7	1.04	(0.89, 1.11)	0.8
2	Motion.engi..	-2.279	0.519	0.79	(0.72, 1.28)	-1.5	0.97	(0.10, 1.90)	0.1
3	Motion.slow..	0.482	0.221	1.17	(0.72, 1.28)	1.2	1.10	(0.87, 1.13)	1.6
4	Motion.spee..	-0.566	0.269	0.88	(0.72, 1.28)	-0.8	0.95	(0.69, 1.31)	-0.2
5	energy.ener..	2.055	0.234	1.09	(0.72, 1.28)	0.6	1.03	(0.79, 1.21)	0.3
6	energy.engi..	2.222	0.242	0.96	(0.72, 1.28)	-0.2	0.98	(0.76, 1.24)	-0.1
7	energy.engi..	-0.944	0.302	0.98	(0.72, 1.28)	-0.1	0.98	(0.59, 1.41)	-0.0
8	energy.ener..	-0.615	0.273	0.87	(0.72, 1.28)	-0.9	0.91	(0.67, 1.33)	-0.5
9	idKinetic.k..	-2.966	0.762	0.44	(0.72, 1.28)	-5.0	0.93	(0.00, 2.31)	0.1
10	idKinetic.p..	-0.691	0.279	1.16	(0.72, 1.28)	1.1	1.04	(0.65, 1.35)	0.3
11	predictKine..	-0.163	0.244	0.93	(0.72, 1.28)	-0.5	0.96	(0.77, 1.23)	-0.3
12	kineticChoo..	2.998	0.301	0.96	(0.72, 1.28)	-0.2	1.03	(0.58, 1.42)	0.2
13	kineticChoo..	2.747	0.278	0.98	(0.72, 1.28)	-0.1	1.04	(0.64, 1.36)	0.3
14	testKinetic..	1.173	0.216	0.98	(0.72, 1.28)	-0.1	0.99	(0.90, 1.10)	-0.2
15	testKinetic..	1.307	0.217	1.12	(0.72, 1.28)	0.9	1.12	(0.89, 1.11)	2.1
16	evaluateKin..	0.196	0.231	0.97	(0.72, 1.28)	-0.1	0.98	(0.83, 1.17)	-0.2
17	explainKine..	1.270	0.214	1.10	(0.72, 1.28)	0.7	1.08	(0.90, 1.10)	1.6
18	explainKine..	-0.626	0.273	0.94	(0.72, 1.28)	-0.4	0.96	(0.67, 1.33)	-0.2
19	explainKine..	-1.379	0.351	0.75	(0.72, 1.28)	-1.9	0.91	(0.46, 1.54)	-0.2
20	explainKine..	1.627	0.219	1.08	(0.72, 1.28)	0.6	1.05	(0.86, 1.14)	0.7
21	predictColl..	-0.733	0.279	0.94	(0.72, 1.28)	-0.4	0.97	(0.65, 1.35)	-0.1
22	predictColl..	-1.080	0.312	0.84	(0.72, 1.28)	-1.1	0.93	(0.56, 1.44)	-0.2
23	predictColl..	2.930	0.292	1.67	(0.72, 1.28)	4.0	1.17	(0.60, 1.40)	0.9
24	Collision.a..	-2.022	0.463	0.61	(0.72, 1.28)	-3.2	0.92	(0.21, 1.79)	-0.1
25	Collision.a..	-0.865	0.293	0.78	(0.72, 1.28)	-1.7	0.92	(0.61, 1.39)	-0.3
26	evaluateTot..	0.017	0.234	1.01	(0.72, 1.28)	0.1	1.02	(0.80, 1.20)	0.2
27	explainTruc..	-2.979	0.764	1.48	(0.72, 1.28)	3.0	1.01	(0.00, 2.32)	0.2
28	explainTruc..	-1.151	0.323	1.07	(0.72, 1.28)	0.6	0.97	(0.53, 1.47)	-0.0
29	explainTruc..	-2.025	0.464	0.77	(0.72, 1.28)	-1.8	0.93	(0.21, 1.79)	-0.1
30	explainTota..	-0.486	0.262	0.79	(0.72, 1.28)	-1.5	0.90	(0.70, 1.30)	-0.6
31	explainTota..	0.708	0.215	0.91	(0.72, 1.28)	-0.6	0.92	(0.89, 1.11)	-1.4
32	evidenceTra..	-1.262	0.336	0.87	(0.72, 1.28)	-0.9	0.96	(0.50, 1.50)	-0.1
33	evidenceDes..	-0.037	0.236	1.19	(0.72, 1.28)	1.3	1.09	(0.79, 1.21)	0.9
34	evidenceDes..	0.218	0.226	1.07	(0.72, 1.28)	0.6	1.02	(0.83, 1.17)	0.2
35	collisionTe..	0.314	0.223	0.89	(0.72, 1.28)	-0.8	0.91	(0.85, 1.15)	-1.1
36	collisionTe..	-2.298	0.520	0.95	(0.72, 1.28)	-0.3	0.99	(0.10, 1.90)	0.1
37	collisionTe..	0.266	0.225	0.88	(0.72, 1.28)	-0.8	0.91	(0.84, 1.16)	-1.1
38	collisionCo..	0.166	0.230	1.08	(0.72, 1.28)	0.6	1.04	(0.82, 1.18)	0.4
39	collisionCo..	0.457	0.221	0.88	(0.72, 1.28)	-0.8	0.91	(0.87, 1.13)	-1.4
40	collisionRe..	0.119	0.230	0.88	(0.72, 1.28)	-0.9	0.91	(0.82, 1.18)	-1.0
41	collisionRe..	0.169	0.228	0.88	(0.72, 1.28)	-0.9	0.91	(0.83, 1.17)	-1.1
42	collisionRe..	-3.027	0.769	0.85	(0.72, 1.28)	-1.0	0.99	(0.00, 2.33)	0.2
43	collisionRe..	0.067	0.232	0.87	(0.72, 1.28)	-0.9	0.90	(0.81, 1.19)	-1.0
44	Intro.poten..	-0.785	0.286	0.64	(0.72, 1.28)	-2.9	0.84	(0.63, 1.37)	-0.9
45	Intro.kinet..	0.198	0.228	0.92	(0.72, 1.28)	-0.5	0.94	(0.83, 1.17)	-0.7
46	ChooseGraph..	0.248	0.226	0.92	(0.72, 1.28)	-0.6	0.95	(0.84, 1.16)	-0.5
47	Predict.top..	-0.623	0.273	0.98	(0.72, 1.28)	-0.1	0.99	(0.67, 1.33)	-0.0



SimScientists WestEd

COPYRIGHT WESTED, 2013

Embedded Assessment 1
Excerpt of Bayes' Net

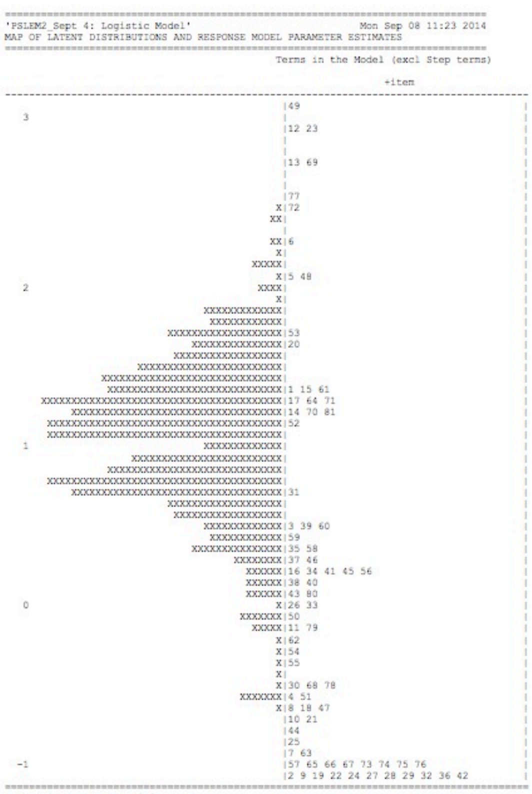


SimScientists WestEd

COPYRIGHT WESTED, 2013

Triangulation

Psychometric Analysis



Cognitive Lab (Think Aloud)

Observations	Reflections
00:26–02:22	Students spent two minutes running the animation twice, and then answering three questions.
PB: So the force is going to be smaller or larger … OK, where do you want to drag the arrow? OA: Right PB: How much OA: 30	Students understand the directions for dragging the arrow.
PB: OK (clicks RUN) That goes faster, right? OA: A lot faster	Students observe the change in speed.
OA: A small forward force causes the truck to slow down. What does a small forward mean? A small force? PB: (laughs) Oh, a small force! … No OA: No (clicks No)	Wording may be confusing because there are two adjectives (small, forward), possibly made more confusing because “forward” can be a noun.
OA: Any forward force causes the truck to speed up. Yeah PB: Yeah, because of the arrows	Students seem to be applying what they observed with different force arrows to the item.

Classroom Observation

8:15–8:20 Teacher introduces activity; students log in.
8:20–8:24 Almost all students have started.
Screen 2: One student compared larger and smaller forces; initially picked yes for smaller force, then changed to no before clicking next.
Screen 3: After reading feedback that says a force transfers energy, one student changes three responses from incorrect to correct.
8:27 A few students can't get started. The Start or Continue button opens an empty window and the assessment doesn't load. These students changed browsers from Firefox to Chrome, which resolves the issue. Around the room, other students are using either Firefox or Chrome successfully.
Screen 5: A student asks why they didn't get correct feedback: “Did I need to write something?”
Screen 7: Students tend to start by creating four trials, some with no variable constant, and some with one variable constant across all four. One student gets feedback about keeping mass constant. This student changes trial 1 so it has the same mass as trial 2, then gets the validation message "You have already saved a trial with these conditions." The student then revises this trial again to vary the speed, and gets feedback about keeping speed constant. The student changes trial 2, then trial 1 again, and then gets correct feedback.

SimScientists WestEd

COPYRIGHT WESTED, 2013

Energy Embedded Assessment Excerpt of Data Mining

Variable Values	[0,0,0,0,0,0,1]	**VariableValuesTotals**	1
Variable Values	[0,0,0,0,1,1,1]	**VariableValuesTotals**	2
Variable Values	[0,0,0,0,1]	**VariableValuesTotals**	1
Variable Values	[0,0,0,1,1,1,1,0,1]	**VariableValuesTotals**	1
Variable Values	[0,0,0,1,1,1,1]	**VariableValuesTotals**	1
Variable Values	[0,0,0,1,1]	**VariableValuesTotals**	2
Variable Values	[0,0,0,1]	**VariableValuesTotals**	4
Variable Values	[0,0,1,1,1]	**VariableValuesTotals**	1
Variable Values	[0,0,1,1]	**VariableValuesTotals**	4
Variable Values	[0,1,0,1]	**VariableValuesTotals**	1
Variable Values	[0,1,1,1,0,1]	**VariableValuesTotals**	1
Variable Values	[0,1,1,1,1,1]	**VariableValuesTotals**	1
Variable Values	[0,1,1,1]	**VariableValuesTotals**	3
Variable Values	[0,1,1]	**VariableValuesTotals**	2
Variable Values	[0,1]	**VariableValuesTotals**	8
Variable Values	[1,0,0,0,0,1]	**VariableValuesTotals**	1
Variable Values	[1,0,0,1,1]	**VariableValuesTotals**	1
Variable Values	[1,0,0,1]	**VariableValuesTotals**	1

Potential energy disappears as kinetic energy is created. ○ Yes ◉ No

Variable Values	[1,0,1,1]	**VariableValuesTotals**	4
Variable Values	[1,1,0,0,0,0,1]	**VariableValuesTotals**	3
Variable Values	[1,1,0,0,1,1]	**VariableValuesTotals**	1
Variable Values	[1,1,0,0,1]	**VariableValuesTotals**	1
Variable Values	[1,1,0,1]	**VariableValuesTotals**	2
Variable Values	[1,1,1,0,0,0,0,1]	**VariableValuesTotals**	1
Variable Values	[1,1,1,1,1,0,1]	**VariableValuesTotals**	1
Variable Values	[1,1,1,1,1,1]	**VariableValuesTotals**	1
Variable Values	[1,1,1,1,1]	**VariableValuesTotals**	2
Variable Values	[1,1,1,1]	**VariableValuesTotals**	1
Variable Values	[1,1,1]	**VariableValuesTotals**	5
Variable Values	[1,1]	**VariableValuesTotals**	13
Variable Values	[1]	**VariableValuesTotals**	26
		Grand Total	98

COPYRIGHT WESTED, 2013

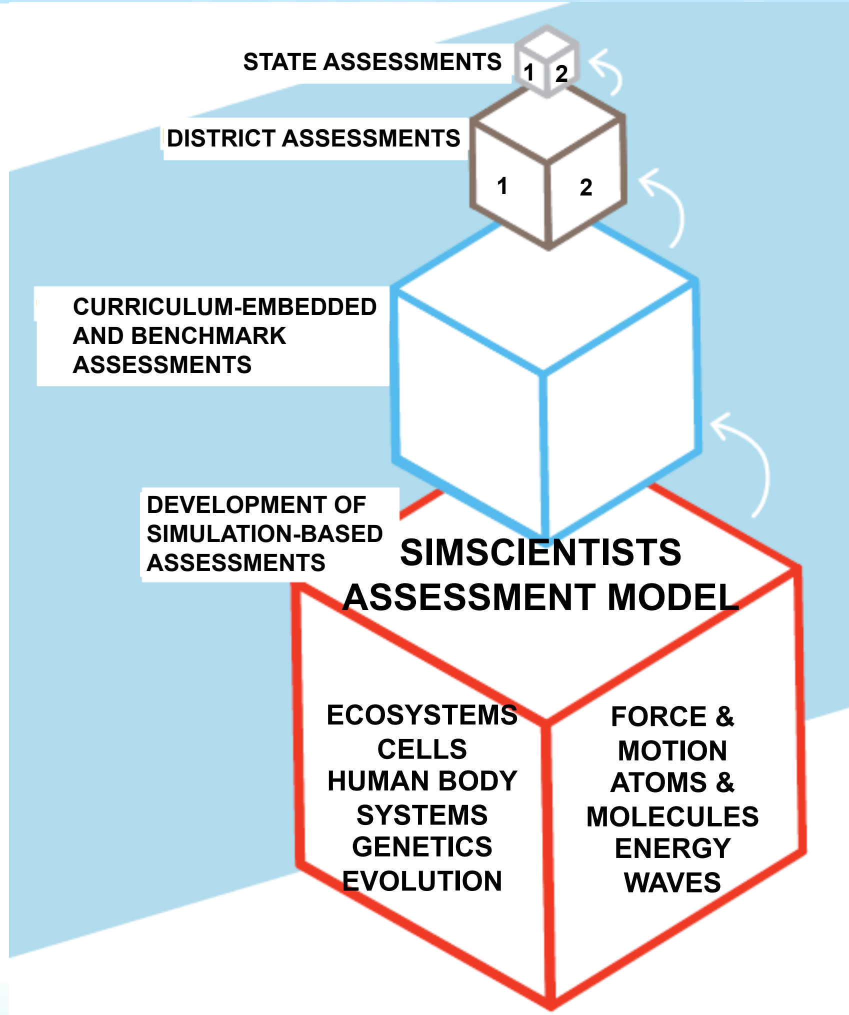
Validation Design

Year 3 Pilot & Validation	Year 4 Cross-Validation
Sample 1: 5 teachers (20 classrooms, 500 students) Use embedded, benchmark, pre/post, signature tasks	Sample 3: 20 teachers (20 classrooms, 500 students) Use embedded, benchmark, pre/post, signature tasks
Sample 2: 5 teachers (20 classrooms, 500 students) Only post test, signature tasks	Sample 4: 20 teachers (20 classrooms, 500 students) Only post test, signature tasks

COPYRIGHT WESTED, 2013

Balanced, Multilevel Assessment System Models

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



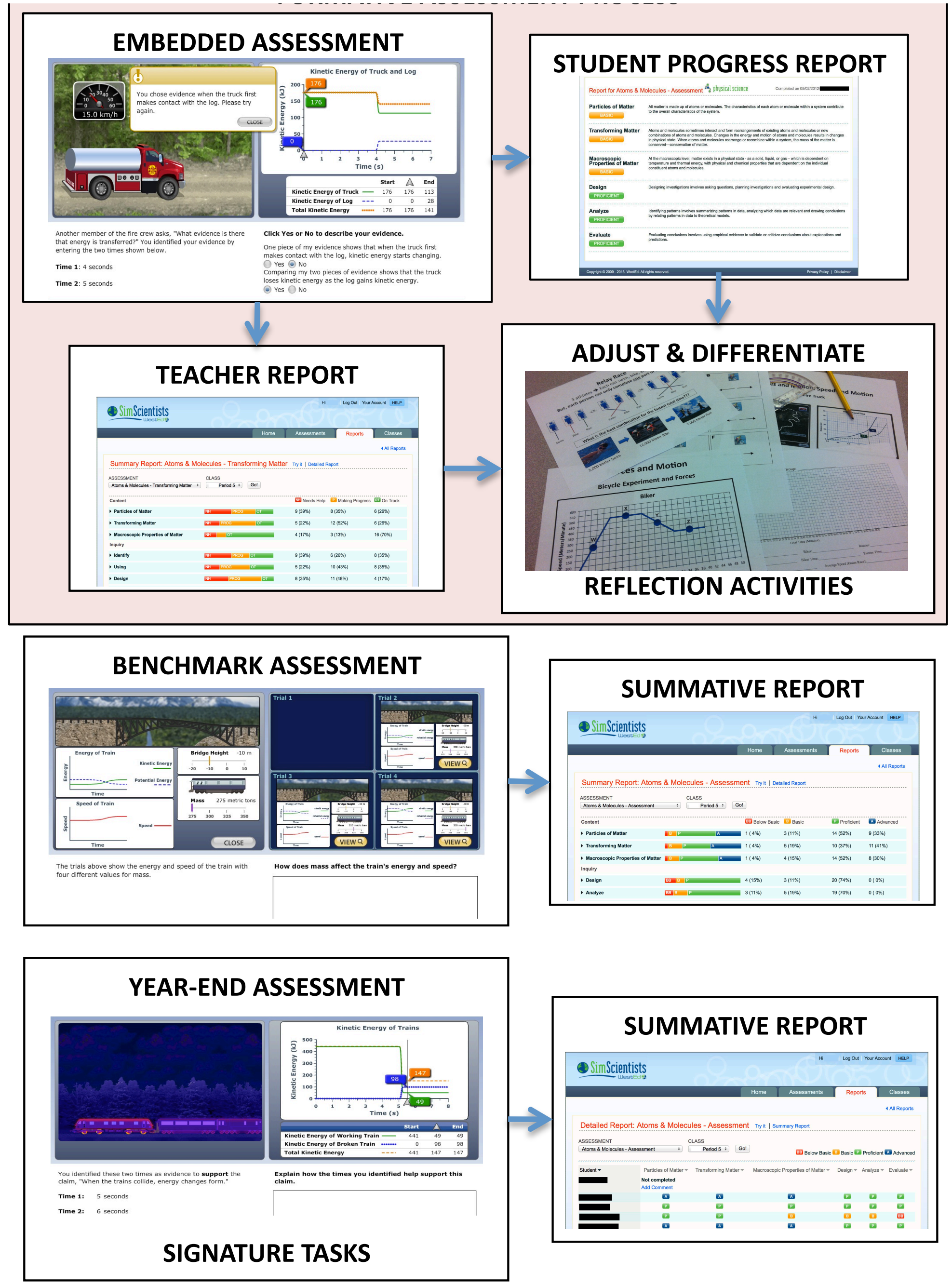
SimScientists WestEd

COPYRIGHT WESTED, 2013

Multilevel Assessment Systems

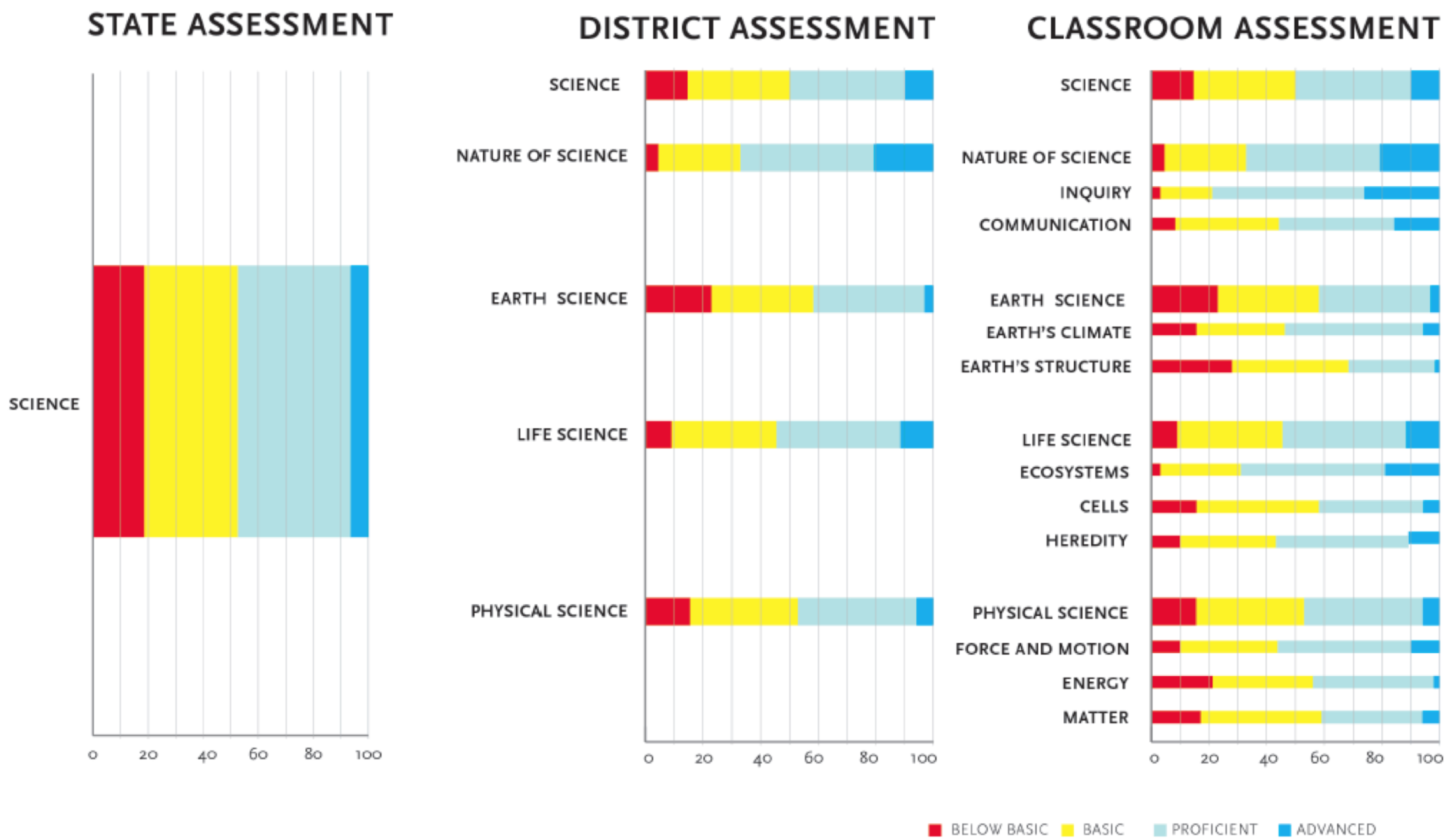
- Integrated assessment design—using common specifications to develop parallel tasks for different levels of the system

- Integrated report design—gathering data from all levels of the system



COPYRIGHT WESTED, 2013

Side-by-Side Model



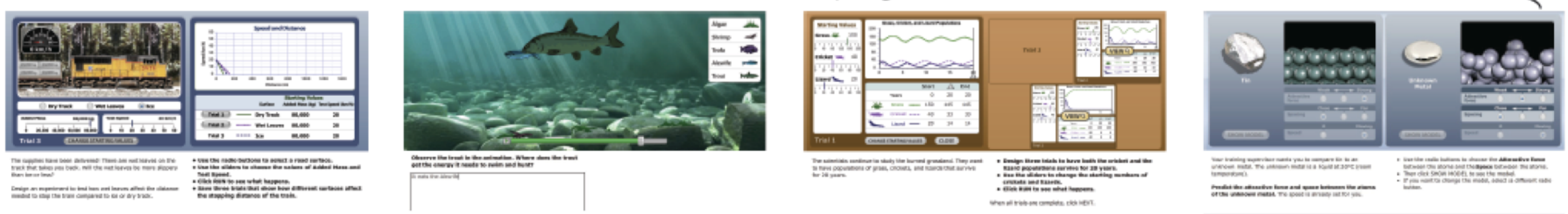
SimScientists WestEd

COPYRIGHT WESTED, 2013

Signature Task Model

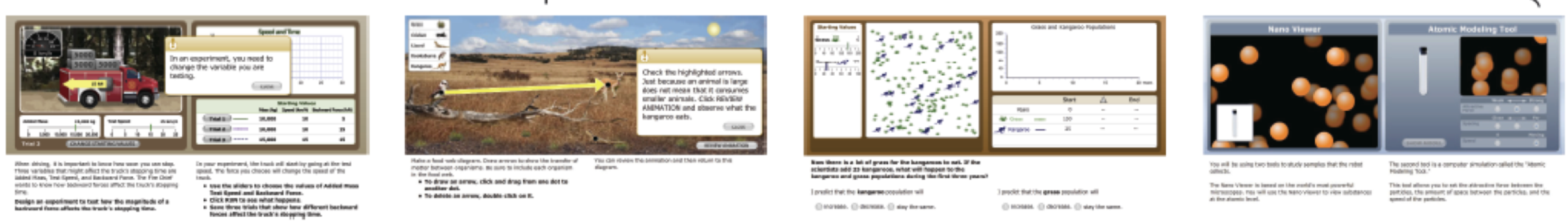
State Test Forms

Matrix Sampling



Simulation-based task item bank

Specifications and Simulation environments



Simulation-Based Classroom Assessments

SimScientists WestEd

COPYRIGHT WESTED, 2013

Research Findings

Impacts on learning

- Two assessment suites in 2 states with 26 teachers and 3,694 students
- Significantly better performance on pre/post and benchmark assessments comparing students using or not using curriculum-embedded assessments

Feasibility of implementation

- Two suites in 3 states, 28 districts, 39 schools, 55 teachers, 5, 465 students
- Can implement across range of infrastructures, teachers, students

Technical quality

- Simulation-based summative benchmark and curriculum-embedded formative assessments

Models for integrating simulation-based assessments into state science assessment systems

COPYRIGHT WESTED, 2013

Selected Publications

Quellmalz, E. S., Davenport, J. L., Timms, M. J., DeBoer, G., Jordan, K., Huang, K., & Buckley, B. (2013). Next-Generation Environments for Assessing and Promoting Complex Science Learning. *Journal of Educational Psychology*.

Buckley, B. C., & Quellmalz, E. S. (2013). Supporting and Assessing Complex Biology Learning with Computer-based Simulations and Representations. In D. Treagust & C.-Y. Tsui (Eds.), Multiple Representations in Biological Education (pp. 247-267). Dordrecht: Springer.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching, 49*(3), 363–393.

Quellmalz, E. S., Timms, M. J., Buckley, B. C., Davenport J., Loveland, M., & Silberglitt, M. D. (2012). 21st century dynamic assessment. In M. Mayrath, J. Clarke-Midura, & D.H. Robinson, (Eds.) *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research.* Information Age.

Quellmalz, E. S., Timms, M. J., & Buckley, B. C. (2010). The promise of simulation-based science assessment: The Calipers project. *International Journal of Learning Technologies, 5*(3), 243–265.

Quellmalz, E. S., Timms, M. J., & Schneider, S. A. (2009). *Assessment of student learning in science simulations and games*. Paper commissioned by the National Academy of Science.

Quellmalz, E. S. & Pellegrino, J. W. (2009). Technology and testing. *Science,* 323, 75–79.

SimScientists WestEd

COPYRIGHT WESTED, 2013

Contact Information

Website:

www.simscientists.org

Email:

equellm@wested.org

msilber@wested.org

COPYRIGHT WESTED, 2013