ONPAR®
Innovative Assessment for Diverse Students

Inclusive Innovations:
Essential Characteristics of Task Design in Determining the Technical Defensibility for Students with Language and Related Challenges

Jesse Markow, WIDA Consortium at WCER, University of Wisconsin
Cathy Cameron, Center for Applied Linguistics

ONPAR®
Innovative Assessment for Diverse Students

Introduction

- State assessments have a number of significant challenges
  - Rigorous content standards have been developed in both English language arts and mathematics, with areas such as science not far behind
  - Race to the Top consortia are moving towards universal online assessment
  - They both want to use a broader array of assessments – more written response and performance assessments
  - Yet, there are a significant number of students who have language challenges – English learners (EL's), students with mild to moderate disabilities, and so forth
- These students have to (and should) participate in the state assessments, but can they participate in a way that permits them to show what they know and can do on academic assessments?
- Will the new state assessment consortia be able to create academic assessments – in mathematics, science, and perhaps language arts – that will prove to be on-grade level in rigor, yet accessible to these students with limited reading and writing capabilities?

ONPAR®
Innovative Assessment for Diverse Students

The Problem

- Some students with normal intelligence cannot show what they know using language and general test formats as the primary communication vehicles.
- These include:
  - Lower English proficient English Learners (ELs)
  - Some students with learning disabilities
  - Other students with disabilities, such as some DHH, ED, ADD, speech, and G & T
  - Poor readers who are not otherwise identified
- At the same time, these students should be exposed to challenging content (and will be assessed on such content in the academic assessments).
- The conundrum is cognitively complex content is often intertwined with language that is too sophisticated.
- However, these students CAN still learn complex content because they and their teachers have learned to convey meaning using other semiotic representations as their primary communication methods.

What Does This Type of Assessment Look Like?

ONPAR®
Innovative Assessment for Diverse Students

- For many of these students, using typical accommodations with most traditional or innovative tasks is not sufficient.

- This is because these students share a difficulty accessing the assessment language or being able to attend to it as it is
  - Used in typical close-ended formats (within traditional or innovative tasks)
  - Scored in general automatic scoring routines of open-response items

- Problems with access occur in how tasks are presented AND in the types of limited response genres within which students must respond.

- This means successful adaptations would need to include accessible ways to both convey meaning:
  - To the student from the test maker
  - From the student to the test maker

ONPAR®
Innovative Assessment for Diverse Students

Defining an Approach

- In order to defensibly integrate these types of tasks into general assessment systems at all levels, evidence about what these tasks measure can be defined by a framework spelling out what meaning is intended and how the tasks convey it.
- While the framework would fit within a more general system's approach, it is needed because the approach to communicating meaning is distinct.
- A more specific framework provides the basis for a detailed content validation crosswalk to other assessment elements in the system.
- Adequate documentation in each section of the framework also provides the basis for interpreting other technical evidence.

ONPAR®
Innovative Assessment for Diverse Students

ONPAR

- ONPAR mitigates language by using deliberate computer-interactive, multi-sensory strategies for
  - Building up and communicating the problem environment and target questions
  - Designing response spaces that allow students to successfully demonstrate their knowledge and skills using a variety of methods
- Instead of using general test formats (e.g., multiple-choice or written response items) ONPAR uses demonstration, continuous interaction, and related creation techniques to primarily convey meaning.
- ONPAR is designed to measure cognitively complex skills and concepts, as well as more basic knowledge which traditional formats cannot assess without a substantial language load.

ONPAR®
Innovative Assessment for Diverse Students

ONPAR Framework

- The ONPAR framework is divided into six environments or sections, each responsible for defining techniques, development strategies, and evidence associated with it:
  - Specification of the Intended Claim(s) of the Task
  - Specification of the Architectural Load of the Task
  - Definition of the Context Environment
  - Construction of the Problem Environment
  - Communication of the Target Question(s) or Statement(s)
  - Specification of the Response Environment(s)
- Each environment focuses on activating particular cognitive processes to convey meaning using relevant representations at specified junctures.
- By designing tasks where meaning is explicitly and strategically introduced, supported, and distributed, complex concepts and skills can be assessed to and from the test taker in novel ways and with little language.

ONPAR®
Innovative Assessment for Diverse Students

Architectural Load of Tasks

- In designing a house or building, based on purpose and function, load-bearing walls are placed strategically in different configurations in order to make the structure stable and strong/defensible.

- The ONPAR design includes considering how to communicate meaning of
  - The general context
  - The specific target

  when little language is given to or collected from the students.

ONPAR®
Innovative Assessment for Diverse Students

Architectural Load of Tasks

- To do this successfully, the communication vehicle and presentation and response formats in ONPAR need to vary between tasks (and from the formats generally used).

- Within each environment they appropriately operationalize the purpose and how it functions.

- Further,
  - Precision of targeted elements needs to be maintained
  - Flexibility in how meaning is conveyed leads to the use of several techniques. These techniques are chosen based on the:
    - Nature of the problem or response
    - Cognitive complexity of the overall task and the particular aspect being operationalized with the technique

- Standardizing elements ground the tasks.

ONPAR®
Innovative Assessment for Diverse Students

Architectural Load of Tasks

- Like evidence-centered design, once claim(s) are identified, item writers
  - Determine what types of evidence from the student is needed to demonstrate their level of the targeted knowledge and/or skills
  - Specify how this evidence can be translated onto the ONPAR screens
- Then, item writers
  - Conceptualize a suitable problem that addresses the claim
  - Determine how the 'meaning loads' need to be distributed across the presentation and response portions of the task
- The tasks MUST successfully communicate without
  - Changing the intent OR cognitive complexity of the task
  - Overwhelming the student
  - Cuing or advantaging the ONPAR test takers when other forms are used with other students in the assessment system to measure the same content

ONPAR Randomized Trials

ONPAR®
Innovative Assessment for Diverse Students

2008 Results in Science

Groups: Lower English-proficient ELs and non-EL Control

Forms: Traditional and ONPAR assessments

1. When controlled for science ability, low English proficient ELs (levels 1 and 2) scored *as well as* non-ELs on the ONPAR tests.
2. Significant differences between ONPAR and traditional forms for low English proficient ELs.
3. NO significant difference between forms for non-ELs.
4. Found similar findings in elementary and middle school.

ONPAR Randomized Trials

ONPAR®
Innovative Assessment for Diverse Students

2009/10 Study in Mathematics

Groups: Lower ELs, students with learning disabilities, students with reading challenges, others

Forms: Traditional and ONPAR

Results to date:

- When controlled for mathematics ability, similar results to science with combination of students with learning disabilities and poor readers and a combination of other students with disabilities.
- Results look similar for low ELs but n so far was too low.

ONPAR®
Innovative Assessment for Diverse Students

Innovative Tasks

- Opening up the response avenues can be defended by linking intended claims and operationalized problems to demonstrate-able solutions.

- Like live performance and performance-based innovative tasks, approaches such as ONPAR can measure complex thinking, including meta-cognitive arguments, inferences, predictions, and complex causal chains in a repeatable administration mode that is also accessible.

- These types of task approaches can be designed to produce summative data, real-time formative information, or both (within constraints of content scope and time).

ONPAR®
Innovative Assessment for Diverse Students

Innovative Tasks

- To a lesser or greater degree, this type of computer-interactive task differs from static language-based task measuring the same content in
  - *Directness* to the latent construct underlying the content target
  - *Response* opportunities
  - *Density* of the cognitive demands (target relevant and avoid irrelevant)
  - *How* target cognitive processes are engaged

Some differences are important and some are not. These need to be reconciled when both are used in assessment systems.

ONPAR®
Innovative Assessment for Diverse Students

Effective Task Design

- The approach also specifies a reasonable framework for specifying how various kinds of innovative tasks might function. For instance,
  - As formative tasks, include a variety of feedback loops. Data from the environments could explain how the processes associated with accessing task aspects changes between loops.
  - For innovative tasks that use Bayesian or other scoring systems, the data points can be linked to techniques and environments activating particular processes.

ONPAR®
Innovative Assessment for Diverse Students

Logistics of Creating Online Interactive Performance Based Assessments

- Delivering ONPAR Online

Cathy Cameron, Center for Applied Linguistics
www.onpar.us

ONPAR®
Innovative Assessment for Diverse Students

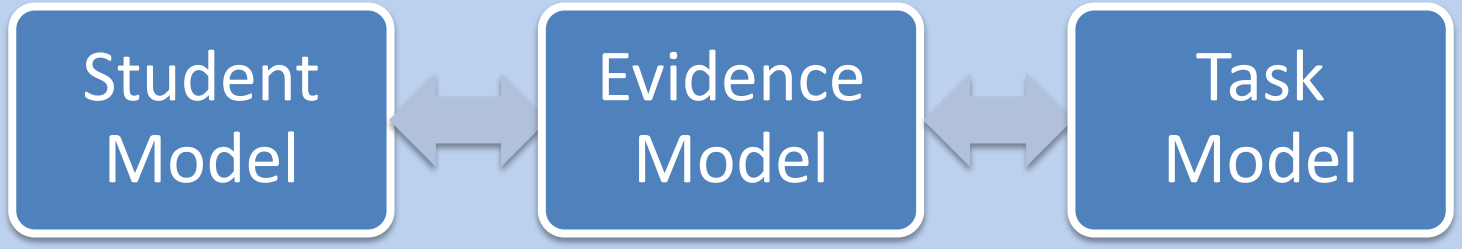
Putting it Together

- Task Development
- Application
  - Provides student interface
  - Incorporates tutorial and tasks
  - May include other elements, such as language selection, questionnaires to collect other student information
- Platform
  - Application presentation tool
  - Includes communication between student and data storage.
  - Incorporates security elements

ONPAR®
Innovative Assessment for Diverse Students

Task Development

- How we ensure we are measuring what we think we are
  - use **Conceptual Assessment Framework** from Evidence Centered Design (Mislevy et al, 1999)



  - Determine what you want to measure
    - consider curriculum, target, expert opinion
  - Determine what kind of task could students do to demonstrate.
    - Evaluate through cognitive labs
  - Design the task that elicits feature

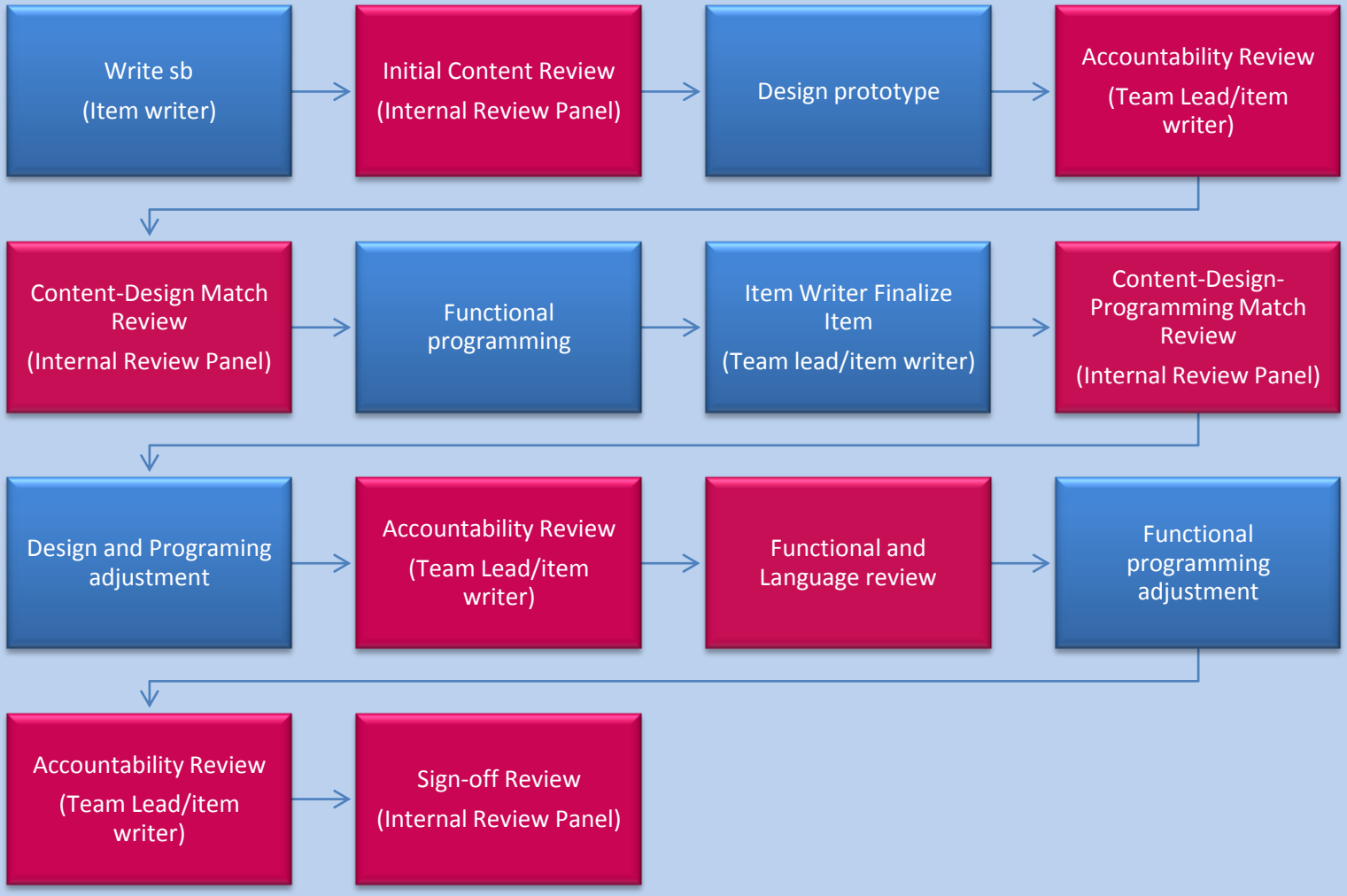
ONPAR®
Innovative Assessment for Diverse Students

Task Development

- **Multisemiotic representations** replace large amounts of text, and **build meaning and complexity across screens**
- **Multisemiotic representations focus students on item target –** provide tools and create meaning
- **Using varied item type –** direct measurement of the construct
- **Intuitive user-interface design –** minimizes the construct irrelevant variance associated with figuring out how to interact with the item

ONPAR®
Innovative Assessment for Diverse Students

Development Cycle



ONPAR®
Innovative Assessment for Diverse Students

Platform

- Robust and redundant
  - Network
  - Data storage systems
- Must accommodate various:
  - Hardware
  - Browsers
  - Levels of network security
- Saves and recalls student responses, including student interactions
- Security Concerns
  - Student authentication
  - Maintain security of student information
  - Maintain security of tasks
  - Assign correct test form

ONPAR®
Innovative Assessment for Diverse Students

Quality Assurance

Task Level Testing	Functionality Exception testing Interactivity
Application Level Testing	All tasks included in application All options
Platform Level Testing	Testing on complete platform including application hosted on website Repeat application level testing Repeat on various platforms, various browsers
Stress Testing	Verify that several applications can be run simultaneously (multiple users /multiple sites) Watch for delay Try in locations with different bandwidth
Database Verification	All information is being saved correctly – answers, scores, interactions Data is being backed up (up to the minute) Query all points of communication
Technology Check	Demo application Minimum Required software/hardware Connectivity

ONPAR®
Innovative Assessment for Diverse Students

Scoring

- All tasks computer scored
- Item writers create rubrics with input from measurement experts and content experts
- Computer programmers convert rubrics to computer algorithms
- Levels of score(partial credit model) based on levels of knowledge that may be demonstrated
- Complex tasks may have multiple correct responses
- When comparing computer to traditional, one rubric for demonstrable knowledge, but algorithm to score will be different.
- Capture interactions
- Analysis from science and math pilots - Scale items for each form (Rasch IRT model to put students and items on the same scale)
- Collect other student data for analysis – demographics, state test scores, ability measure.

ONPAR®
Innovative Assessment for Diverse Students

Sample Size

- Test whole classrooms whenever possible to minimize school disruption – may over sample control
- One cell for each group of interest per form per grade
- Identity the N necessary to attain the desired power to observe the hypothesized effect size for the selected α. from *A Power Primer* (Cohen, 1992)
  - We used about 50/cell for medium effect size for ANCOVA analysis
- Rasch model – closer to 250-500/cell (Embretson & Reise, 2000)

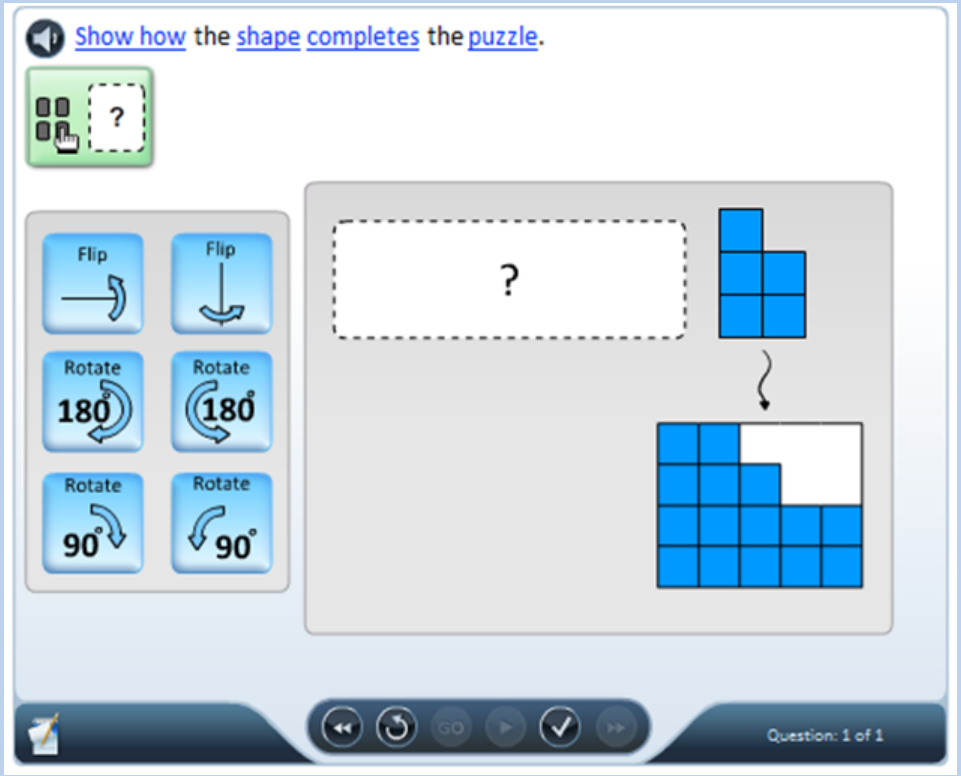
ONPAR®
Innovative Assessment for Diverse Students

Approaching Schools

- Identify the benefit to the school/division/state
- Find a champion within division/state Dept of Ed
- Be generous with compensation
- Be very clear re expectations/compensation
  - Provide specific details re: timeframe, expectations of school/teacher/student, training schedule, sample size, technical requirements (headsets, software, etc.), contingencies, support
- Be sensitive to teacher time
- Allow plenty of lead time for
  - Training – must be organized to meet teacher schedules (record sessions)
    - Online training for division/school coordinator/teachers/test administrator
    - Supporting materials – online manual, demo, help desk
  - Technology changes – may be controlled at division
- Maintain communication/Single point of contact
- Thoroughly test technology – zero tolerance
- Stick with dates – zero tolerance
- IRB in place before contacting
- Ensure security and anonymity of students.

ONPAR®
Innovative Assessment for Diverse Students

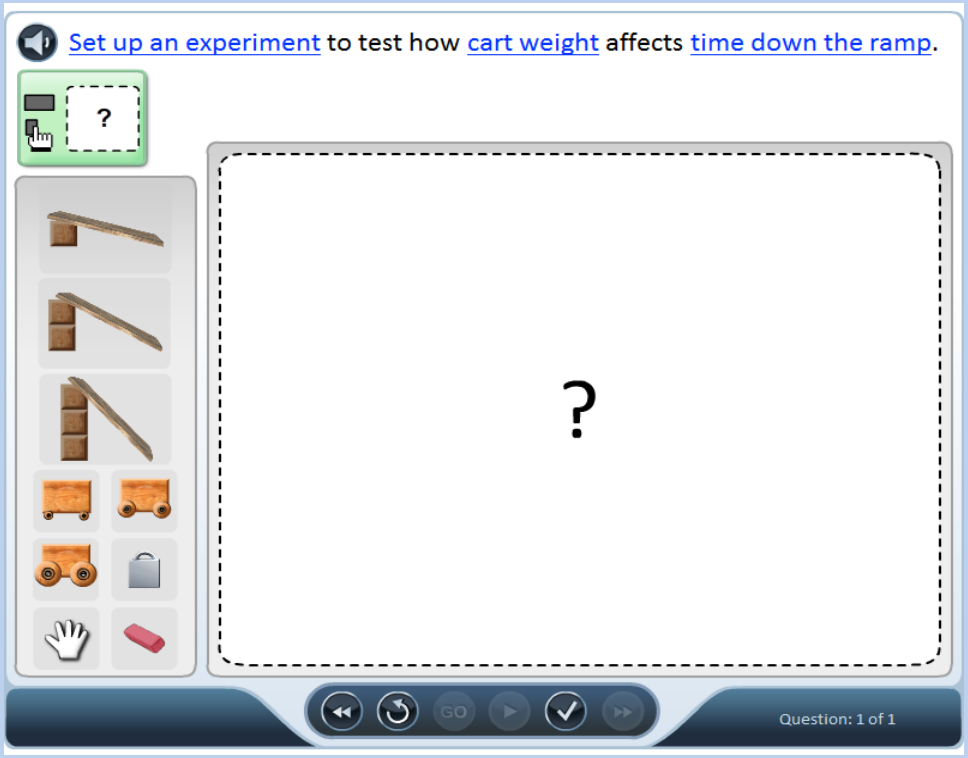
Puzzle



- Animations –apply concepts and then serve as tools
- Build in complexity, can be even more sophisticated
- Producing vs. selecting
- Intuitive user interface design elements

ONPAR®
Innovative Assessment for Diverse Students

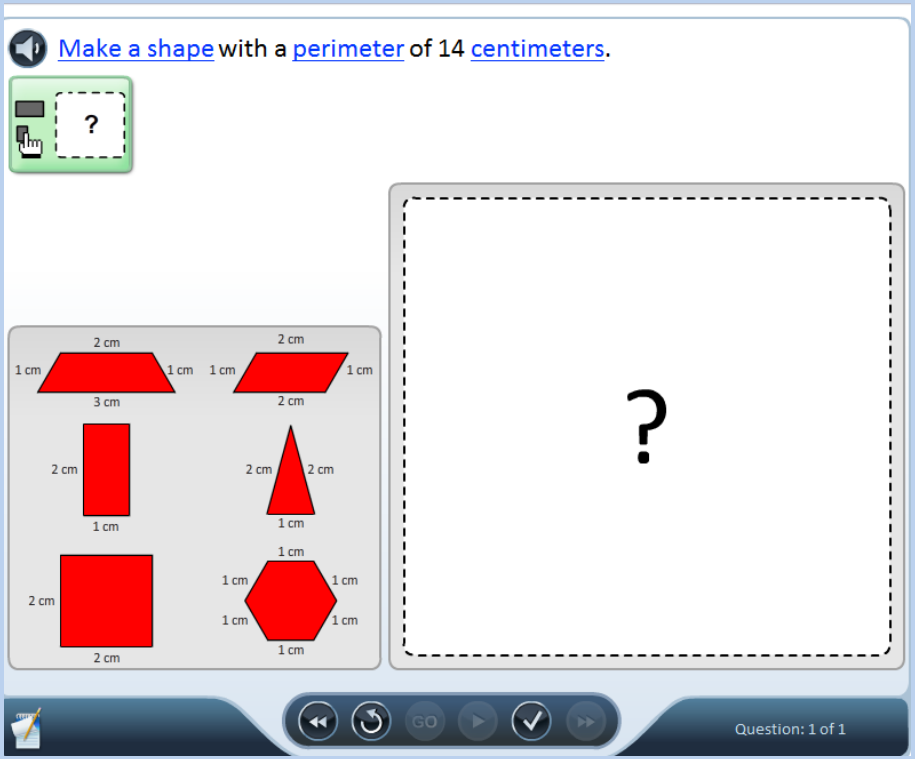
Ramp Experiment



- Building what you are supposed to be thinking about – introduces experiment and variables
- More direct information than selecting or writing
- Intuitive user interface design

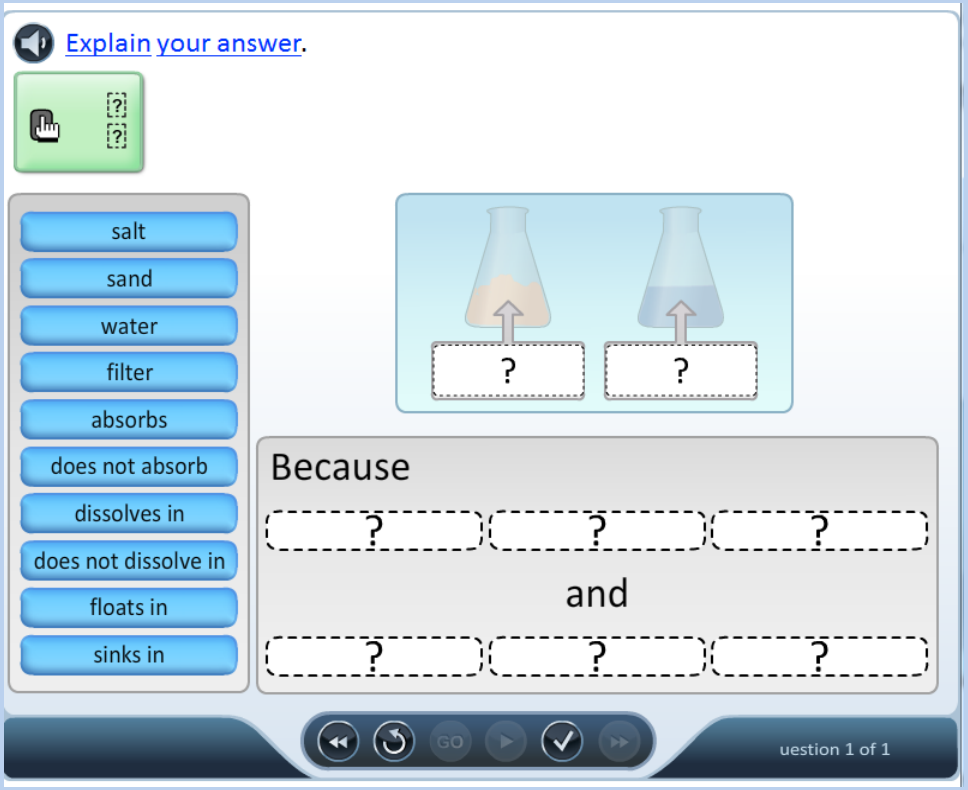
ONPAR®
Innovative Assessment for Diverse Students

Build a Shape



- Simulates building activity
- Gets at the same idea from different angles to measure student knowledge
- More direct information than selecting or writing

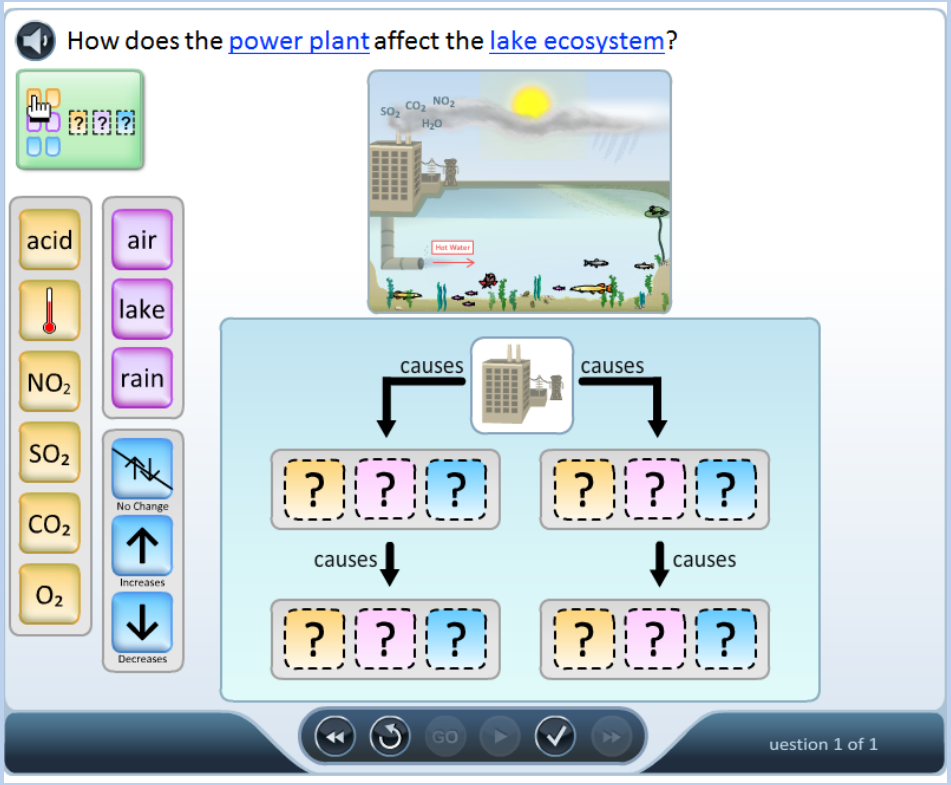
Substance Separation



- Builds concepts and complexity through several question screens to the target task question
- Animations – essential, rather than on the periphery or for engagement purposes
- Continue to support but need for the support decreases

ONPAR®
Innovative Assessment for Diverse Students

Power Plant



- Response environment can match the complexity of the construct and target

ONPAR®
Innovative Assessment for Diverse Students

- Thank You!

- www.onpar.us