Redesigning testing in science

Bringing research-based diagnostic assessments into the classroom

A workshop for NSTA
April 13, 2013
Overview

- DiscoTest design principles
- Sort task challenge
- Insights from the sciences of learning
- What we’re learning
- Break
- DiscoTests in the classroom
- Coding the POE001
The DiscoTest Vision
DiscoTest: top 10 design principles

1. **evidence-based**—backed by solid research (NRC, 2001, 2006)

2. **relevant**—are direct tests of key knowledge and life skills, not test-taking skills (Frederiksen & Collins, 1989; NSF, 2008; Gordon Commission, 2013)

3. **formative**—are rich learning experiences for teachers and students (Frederiksen & Collins, 1989; Collins & Halverson, 2009)

4. **diagnostic**—determine what each student can do right now, and what comes next. (Frederiksen & Collins, 1989; NRC, 2006, Collins & Halverson, 2009; Gordon Commission, 2013)

5. **embeddable**—are part of the lesson (Collins & Halverson, 2009)
DiscoTest: top 10 design principles

6. **support teaching**—are built with teachers and support professional development (NRC, 2001, 2006; Gordon Commission, 2013)

7. **build knowledge**—contribute to the learning sciences (NRC, 2001, 2006; Gordon Commission, 2013)

8. **low stakes**—lots of tests, lots of topics, no test anxiety (NRC, 1999, 2006)

9. **standardized**—to a universal learning scale (Dawson & Stein, 2008; President’s Advisory Council, 2012; Gordon Commission, 2013)

10. **broadly available**—leverage technology, serve the least advantaged (NSF, 2008)
Virtuous cycles of learning
Learning is a kind of **virtuous** cycle

- The idea goes back to the first biologically-oriented psychologists
- As an orienting generalization, the idea covers a wide range of important learning theories, especially those falling under the banner of **constructivism**.
The circular reaction: learning at its core

- Baldwin, Piaget, and a baby in a highchair
  - think you got it...
  - try it...
  - get feedback....
  - try again....

- Virtuous cycles always include tests (often failed or only marginally successful)
Learning to walk

No cycles here, just steady progress. But what is it really like?
Learning to walk: tests & feedback

Babies fall down. A lot. And it usually hurts.
Learning to walk: Support

This is called cruising.

This is called getting help.
The virtuous cycle
Testing is a key component
Carrots, scaffolds, practice, & reflection

1. Set *provisional* learning goals

2. Find out what individual learners know and how they think with what they know.

3. Target instruction to the needs of individual learners. Provide learning challenges that are in their *zone*—just beyond their current level of understanding.

   4. Provide ample scaffolding, and help students develop self-scaffolding skills.

5. Always pair instruction with opportunities to apply new knowledge and skills in hypothetical and real-life situations.

6. Cultivate a habit of reflection: require learners to reflect regularly about outcomes associated with the application of new knowledge.

7. Provide timely feedback, focusing on growth over time rather than success or failure.
Top 7 impediments to VCL (and creativity)


2. Teach students to dislike learning (Batey and Furnham, 2006)

3. Ask questions with right answers (Paul, & Elder, 2006)

4. Restrict opportunities for collaborative hands-on learning (Ball, 1994; Hargreaves, 1994; Grant & Berry, 2011; Nickerson, 1999)

5. Support homogeneity (Ekvall, 1991; Rubenson & Runco, 1995; Maddux, Adam, & Galinsky, 2010)

6. Value breadth over depth (Schwartz, Sadler, Sonnert, & Tai, 2009)

7. Value a limited range of skills, knowledge, and modes of learning (Messick, 1976)
The metric
Card sort challenge
The Lectical® Assessment System

- The common core metric behind all DiscoTests
- Based on Fischer’s Skill Scale (1980, 2006)
  - tiers, levels, and phases
  - 4-5 tiers, 14 levels (3 per tier, 0 = birth and 13 = Einstein), 4 phases in each level (a,b,c,d)
- Domain general, content independent (Dawson, 2004; 2011)
  - Unique and powerful psychometric properties (Dawson-Tunik, Commons, Wilson, & Fischer, 2005)
The developmental spiral

Sensorimotor Actions

Representations

Abstractions
No concept of energy, *per se*.

The ball will fall down.
Energy equals motion, or is a feeling you have.

The ball will fall down

so

it will bounce back up

because

it will go really fast.
Energy is present while something moves.

The ball will run out of energy (stop moving)

- it is going really fast
- it gets slower and slower
- it bounces pretty high
- it can't come up too high

So but

So but
Energy causes motion and can move between objects.

The ball will fall down

The ball will get more and more energy while it falls

the ball will go faster and faster

faster-moving things have more energy

things always fall when you drop them
Energy exists in different forms and amounts

The ball will fall down
so
it will bounce back up
because
it will go really fast.

energy increases as things fall
but
friction changes some of the energy into heat
so
when a ball falls, energy changes from one form to another
Energy is conserved.

The total energy of the system will be conserved.

When things fall, energy changes from one form to another.

Before it falls, the ball has gravitational potential energy.

The potential energy converts to kinetic energy.

Some kinetic energy converts to heat energy.

Before it falls, the ball has gravitational potential energy.

When things fall, energy changes from one form to another.

The total energy of the system will be conserved.

Energy is conserved.

The ball will run out of energy (stop moving).

It can’t come up too high.

It gets slower and slower.

It bounces pretty high.

It is going really fast.

The ball will run out of energy (stop moving).

But so

it gets slower and slower.

But so

it bounces pretty high.

It can’t come up too high.
What we’re learning
Learning sequences
Learning pathways
## From telling to persuading

<table>
<thead>
<tr>
<th>Personal opinion</th>
<th>Persuasion</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>09a (G 3-4)</strong> It is bad for children to learn bad things. (G time 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>09b (G 4-5)</strong> To convince the other side, tell them [something specific {one 09b or multiple 09a}] that you think or know. (A,C time 1,F,G time 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>09c (G 5-6)</strong> To convince the other side, tell them [something specific {one 09c or multiple 09b} things] that you think or know. (B time 1)</td>
<td>To convince the other side, you could do (or say) [something specific] that will appeal to their feelings or preferences. (A,F,G time 3)</td>
<td>To convince the other side, you can show results from a study [involving one unipolar variable] or other {09c} evidence. (C,D,E)</td>
</tr>
<tr>
<td><strong>09d (G 7-8)</strong> To convince the other side, you can give good examples or reasons (in general). (B time 2,D time 2)</td>
<td>To convince the other side, you can show results from research [involving one bipolar variable] or other {09d} evidence. (C,E)</td>
<td></td>
</tr>
<tr>
<td><strong>10a (G 9-10)</strong> To convince the other side, you can make sure you get your point across. (D)</td>
<td>To convince the other side, you can show results from research [involving two bipolar variables] or other {10a} evidence. (D)</td>
<td></td>
</tr>
</tbody>
</table>
## From telling to evidence

<table>
<thead>
<tr>
<th></th>
<th><strong>Personal opinion</strong></th>
<th><strong>Persuasion</strong></th>
<th><strong>Evidence</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>09a (G 3-4)</td>
<td>It is bad for children to learn bad things. (G)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09b (G 4-5)</td>
<td>To convince the other side, tell them [something specific {one 09b or multiple 09a}] that you think or know. (A,C time 1,F,G)</td>
<td>To convince the other side, you could do (or say) [something specific] that will [appeal to their feelings or preferences]. (A,F,G)</td>
<td>To convince the other side, you can show results from a study [involving one unipolar variable] or other {09c}evidence. (C time 2,D,E)</td>
</tr>
<tr>
<td>09c (G 5-6)</td>
<td>To convince the other side, tell them [something specific {one 09c or multiple 09b} things] that you think or know. (B)</td>
<td>To convince the other side, you can give good examples or reasons (in general). (B,D time 2)</td>
<td>To convince the other side, you can show results from research [involving one bipolar variable] or other {09d}evidence. (C time 3,E)</td>
</tr>
<tr>
<td>09d (G 7-8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10a (G 9-10)</td>
<td>To convince the other side, you can make sure you get your point across. (D time 3)</td>
<td>To convince the other side, you can show results from research [involving two bipolar variables] or other {10a}evidence. (D time 3)</td>
<td></td>
</tr>
</tbody>
</table>
Average Lectical score by grade
compared to Northeastern low and high-SES schools
402 people, 626 DiscoTests

Grade

Average Lectical score

- Your results
- Northeastern low-SES
- Northeastern high-SES
Projected growth comparisons for low and high SES schools
Let’s take a break

› Be back in 15 mins...
Let’s take a break

▶ Be back in 15 mins...
Overview of the second half

- DiscoTests in the classroom
  - assessments, reports, and other resources
- Hands-on experience with POE001 coding
- Concluding thoughts
DiscoTests in the classroom
Introduction

DiscoTest is a new kind of standardized assessment infrastructure. Every assessment event:

- provides a learning experience for students and teachers
- equips teachers with diagnostics for making evidence-based instructional decisions

Every DiscoTest includes:

- the assessment itself (called a Teaser)
- a set of low-inference scoring rubrics
- diagnostic reports
- instructional resources that are aligned with assessment outcomes and curricular objectives
Teasers

- Teasers pose questions that require students to connect ideas, think through problems, and show their reasoning.
Constructed responses

- Students write short “essays” (usually 2 or 3 sentences)
- All questions require students to provide justifications for their judgments
  - items are **real life** and **ill-structured** (no “right” answers, just increasingly adequate answers)
  - students must employ essential writing, reasoning, and argumentation skills to address them
- Responses are coded with low-inference rubrics
Low-inference rubrics

- Teachers (and often students) code Teaser questions with a set of low-inference rubrics.
- Teachers learn to code by calibrating with Lectica scorers.
- Each rubric targets a specific concept or skill.
Low-inference rubrics, cont.

- Each rubric is presented as a developmentally-ordered pull-down menu.
- Teachers select the lowest item in the list that is a good match to what a student has written.
- There are about 20 low-inference rubrics per DiscoTest; usually, 5–10 apply to a given performance.
- Selections are analyzed in real time to calculate a performance’s Lectical level. The score is an index of reasoning complexity (on targeted constructs).
- Teachers may also enter personal feedback.
**Question 1:** The image shows a hand pressing down on a ball. What is happening to the energy of the ball as it is pressed into the spring? Explain your answer.

1(a) Describe what is happening to the energy of the ball as it is pressed into the spring? Explain your answer.

The ball only has energy due to its height from the table, if dropped off the spring. The spring has the energy of compression, unless the ball itself is compressed, as in being of rubber. Nothing would happen to the energy of the ball. The potential energy of the spring would be turned into kinetic energy.

1(b) What would happen to the energy of the ball if the hand moved away? Explain your answer.

---

Score for question 1 10.2
Argumentation scales

- DiscoTest’s low-inference rubrics are designed to address the **complexity** of students’ thinking
- Coders are expected to ignore the **quality** of students’ writing when using the low-inference rubrics
- Writing quality is evaluated with **argumentation** scales
- These address writing **mechanics**, **clarity**, and **coherence**
- Student reports include feedback about both complexity and argumentation
- Today, we will not be using the argumentation scales
Report card and reports

Marilyn Monroe

Lectical score

Grades

Social studies
Physical science
Learning resources

- Learning sequences
- QuickPick Activities
- Coding tutorials (coming soon)
- Suggestions for using DiscoTest diagnostics and resources to customize instruction (coming soon)
- Video clips of exemplary instructional strategies in action (coming soon)
Lesson planning (coming soon)

‣ Vetted, customizable, ready-to-use core lesson plans

‣ Lesson materials and resources

‣ Supplementary materials (e.g. materials for struggling readers, materials for English language learners, additional homework options, extension activities)

‣ Information on how lessons and Teasers align with Common Core State Standards and National subject area standards
Teacher forum (coming soon)

- A rich interactive space in which teachers can build knowledge and proficiency by working with others who are engaged in using DiscoTests
  - Discussion threads, live topic talks, and scheduled Q/A sessions
  - User profiles and teaching biographies, including “Featured Practitioner”
  - Submission area for suggested Quick-Pick activities & lesson plans
  - Lesson and Teaser feedback (from you to us)
  - Video clips of lessons
  - A reward system that will allow teachers to earn points toward “Master Teacher” recognition when they provide feedback, submit quality lessons, or participate in other ways
The POE001: the physics of energy
Targeted constructs

- The POE001 focuses on students’ understanding of **forces, forms of energy, energy transfer, energy conservation, and work**.

- These concepts play an important role in laying the foundation for success in most STEM disciplines (National Science Standards, 1996; Next Generation Science Standards, 2013).
1. (a) Describe what is happening to the energy of the ball as it is pressed into the spring? Explain your answer. (b) What would happen to the energy of the ball if the hand moved away? Explain your answer.

2. Describe what is happening to the energy of the ball as it sits on top of the spring. Explain your answer.

3. Describe what is happening to the energy of the ball as it falls to the floor. Explain your answer.

4. Describe what is happening to the energy of the ball as it hits the floor. Explain your answer.

5. Describe what is happening to the energy of the ball right after it hits the floor. Explain your answer.
Coding the POE001
Coding rules

1. aim to match student responses with all rubric choices that apply

2. choose the rubric at the highest level (lowest on the list) if more than one level in a rubric matches a given response

3. aim to infer as little as possible; code only what a student actually says

4. use justification ‘99’ if student gets a concept “backwards”
Discussion
What, even more work?

- How long does it take to learn?
- How long does it take to code?
- Lots of bang for your buck.
  - professional development (credits)
  - rich diagnostics
  - lesson plans
  - Quick-Pick activities
  - increased potential for parental involvement
  - students learn
Other ways to get involved

- Become a professional coder—or just come play in our sandbox

- Become a DiscoTest Network teacher or school
  - Contribute to curriculum development
  - Become a DiscoTest master teacher
  - Become a DiscoTest volunteer ambassador
  - Participate in test development
Concluding thoughts

- Teachers, technology, and the learning sciences will shape the future of testing
- The DiscoTest Initiative is on the cutting edge of these trends
- Get involved
  - your classroom, your school, your research project
  - please see the information provided on your flash drive
- Thank you!
Quick-Pick Activities

- Quick-Pick Activities are enrichment resources that are tailored for a student’s current level of performance.
- They are included in Teaser reports, DiscoTest report cards, lesson plans, and resource menus.
- They include:
  - quick experiments
  - questions for independent research
  - activity and media links
  - review activities
  - discussion topics
Formative/summative

- Formative assessments can be embedded in curricula to support learning and provide rich information about student understanding. They often have open-ended constructed-response questions, and incorporate self (or peer) evaluation.

- Summative assessments provide information about the content and procedures students have learned in a given period of time. They primarily consist of selected response (a.k.a. multiple choice) questions and provide little or no formative feedback.