

Assessment & the Common Core: Will our Tests Support or Subvert our Vision of Deeper Learning of English/ Language Arts?

P. David Pearson

UC Berkeley

Slides will be posted at
www.scienceandliteracy.org

My stand on the standards

- Spirit of Critical Colleagueship
 - David C. has listened to my concerns...
- View the CCSS as a work in progress
- Close reading and purpose...
- I desperately want to see them work their magic...
- Moral obligation: Just as we want kids to base their claims on evidence, so we must hold our standards to the same standard.

Assessing Literacy: Problems and Possibilities???

- How did we get to where we are?
- Where are we headed?
- Where should and could we be headed?

P. David Pearson

UC Berkeley

www.scienceandliteracy.org

Some Context for Today's Talk

- My interest stems from several sources
 - Multiple attempts to build state assessment systems
 - Several years of attempts to build alternative assessments—portfolios, performance assessments, formative systems.
 - Attempts to work with schools to build assessment systems rather than collections of tests
- A conviction that accountability should not drive us into a corner in which we engage in practices that are not in the best interests of students, teachers, or the public.
- The opportunity provided by the current forces at work in American education

The Forces

- The swing away from the basics of NCLB
- The swing to the CCSS
- The pressure for deeper learning and 21st century skills...

www.scienceandliteracy.org

NCLB

- No one is fussing as much about the fundamental skills
 - Subword level skills
 - Fluency
- They hover quietly but ever-present in the background..
- A kind of 2 (but not 3) cheers for phonics...

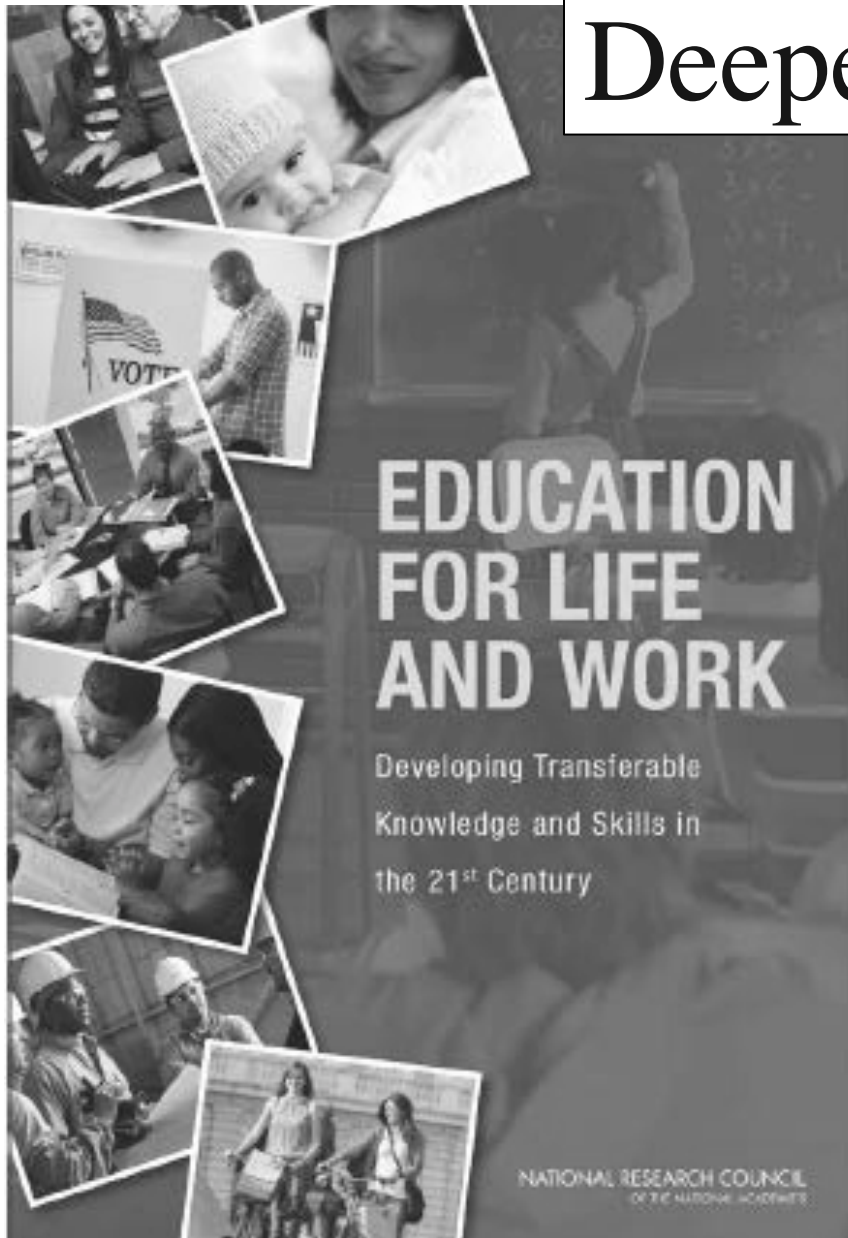
COMMON CORE STATE STANDARDS FOR

English Language Arts
&
Literacy in History/Social Studies,
Science, and Technical Subjects

As a natural outgrowth of meeting the charge to define college and career readiness, the Standards also lay out a vision of what it means to be a literate person in the twenty-first century. Indeed, the skills and understandings students are expected to demonstrate have wide applicability outside the classroom or workplace.

Students who meet the Standards readily undertake the close, attentive reading that is at the heart of understanding and enjoying complex works of literature. They habitually perform the critical reading necessary to pick carefully through the staggering amount of information available today in print and digitally. They actively seek the wide, deep, and thoughtful engagement with high-quality literary and informational texts that builds knowledge, enlarges experience, and broadens worldviews. They reflexively demonstrate the cogent reasoning and use of evidence that is essential to both private deliberation and responsible citizenship in a democratic republic. In short, students who meet the Standards develop the skills in reading, writing, speaking, and listening that are the foundation for any creative and purposeful expression in language.

Deeper Learning



While education has long contributed to national prosperity and stable families and communities, it is more critical than ever in the face of growing environmental, political, and social challenges. To prepare young people to address these challenges, business and education leaders are asking schools to develop skills such as problem solving, critical thinking, communication, collaboration, and self-management —often referred to as “21st century skills” or “deeper learning.”

Assessment History 101

How did we get to where we
are?
Where did it all start?

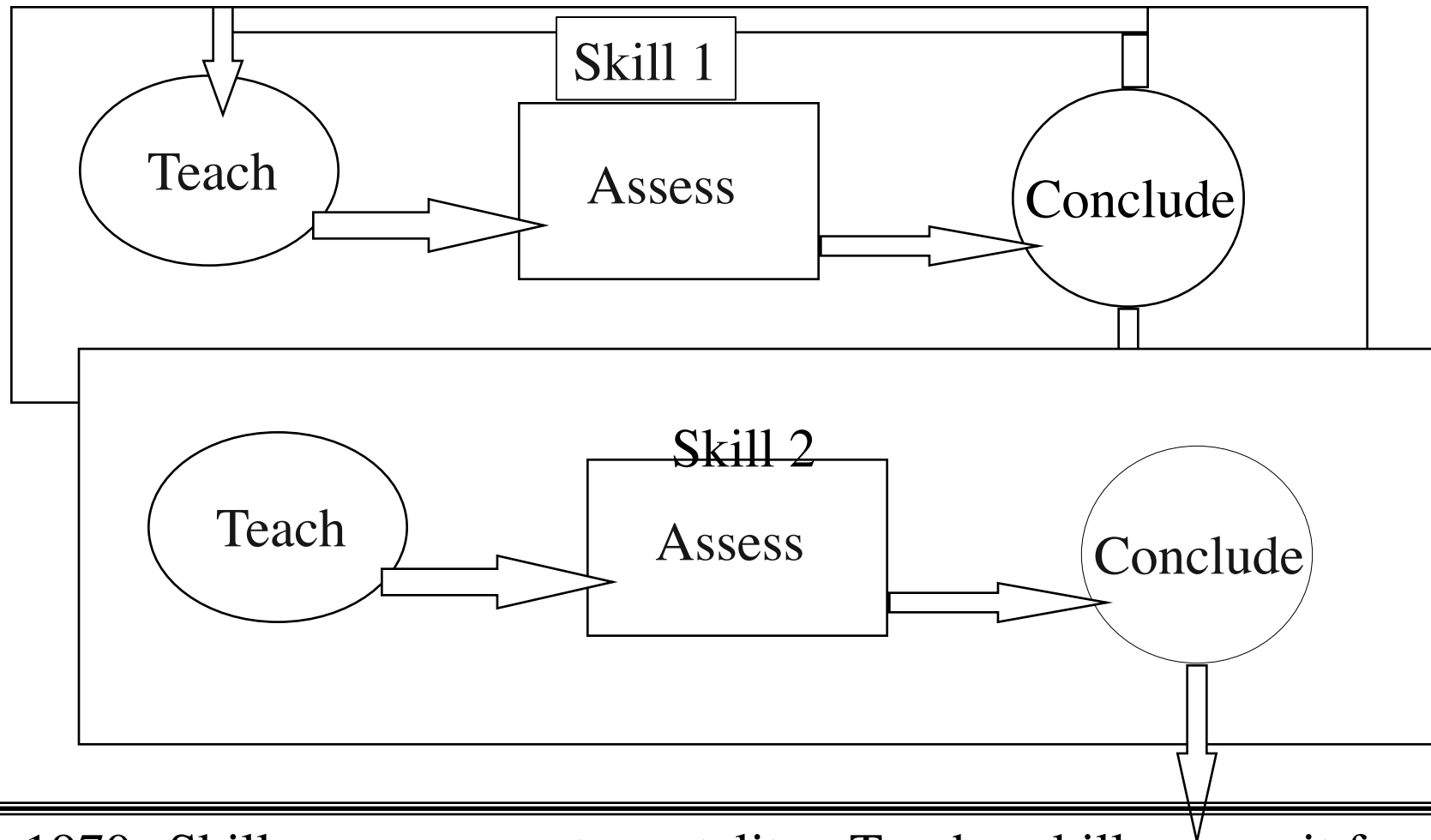
1960s

- Early 1960s: A kinder, gentler time
 - Tests were there
 - Few consequences for anyone—students, teachers, schools, districts
- The turning point:
 - Title 1 reauthorization in 1967: the first trade of accountability of outcomes for flexibility in inputs (curriculum and teaching).

The 1970s

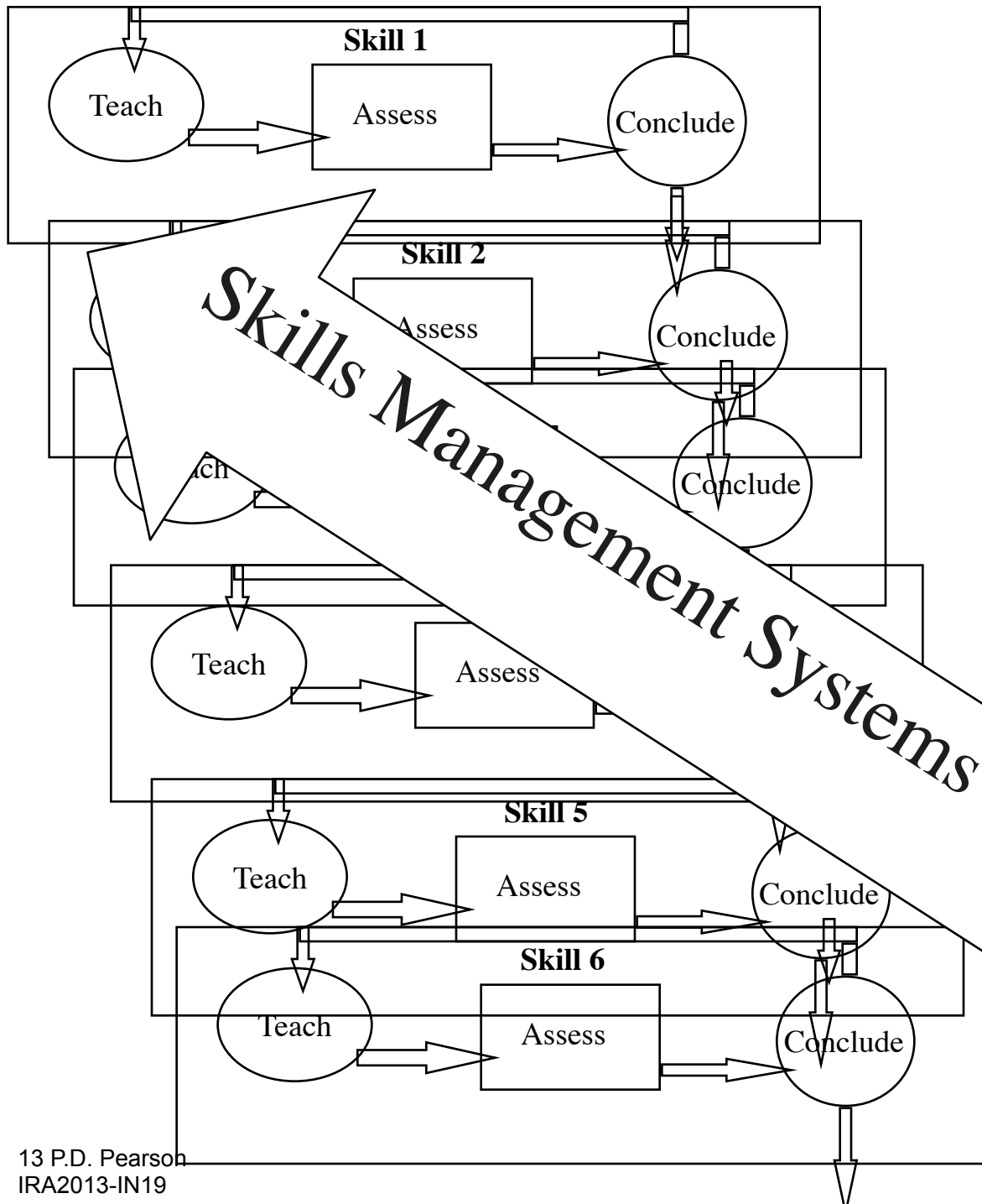
- Behavioral objectives
- Criterion referenced assessments
- Curriculum-embedded assessments
- Minimal competency tests: New Jersey
- Statewide assessments: Michigan & Minnesota

Historical relationships between instruction and assessment



The 1970s Skills management mentality: Teach a skill, assess it for mastery, reteach it if necessary, and then go onto the next skill.

Foundation: Benjamin Bloom's ideas of mastery learning



The 1970s, cont.

And we taught each of these skills until we had covered the entire curriculum for a grade level.

Rule of Thumb

- The finer the grain size at which we monitor a process like reading and writing, the greater the likelihood that we will end up teaching and testing **bits and pieces** rather than **global processes** like comprehension and composition.

The 1980s

- In like a lamb, out like a lion
- Quiet consolidation of skills based learning (especially in commercial materials)
- *1983: A Nation at Risk* leads to call for school reform
- The evolution of “higher order” thinking in all subject matter areas
 - Driven by the business community
- Late 1980s: Do we have assessments to match our models of learning-----
 - Standards: The 1998 Charlottesville Conference
 - Clinton sponsorship
 - Performance and authentic assessment movements
 - Eradicate the distinction between instruction and assessment

Key movements of the late 1980s/1990s

- Kentucky and Vermont portfolio assessments
- Graduation by portfolio presentation
 - Central Park East and Walden
- Assessment by exhibition: Coalition of Essential Schools
- Performance Assessment: New Standards, CLAS, Many States---Maryland, Washington...
- Some models of teacher performance assessment: NBPTS, Connecticut

Late 1980s/early 1990s:
Portfolios
Performance Assessment
Make Assessment Look Like Instruction

Note: We did NOT try to make
instruction look like assessment

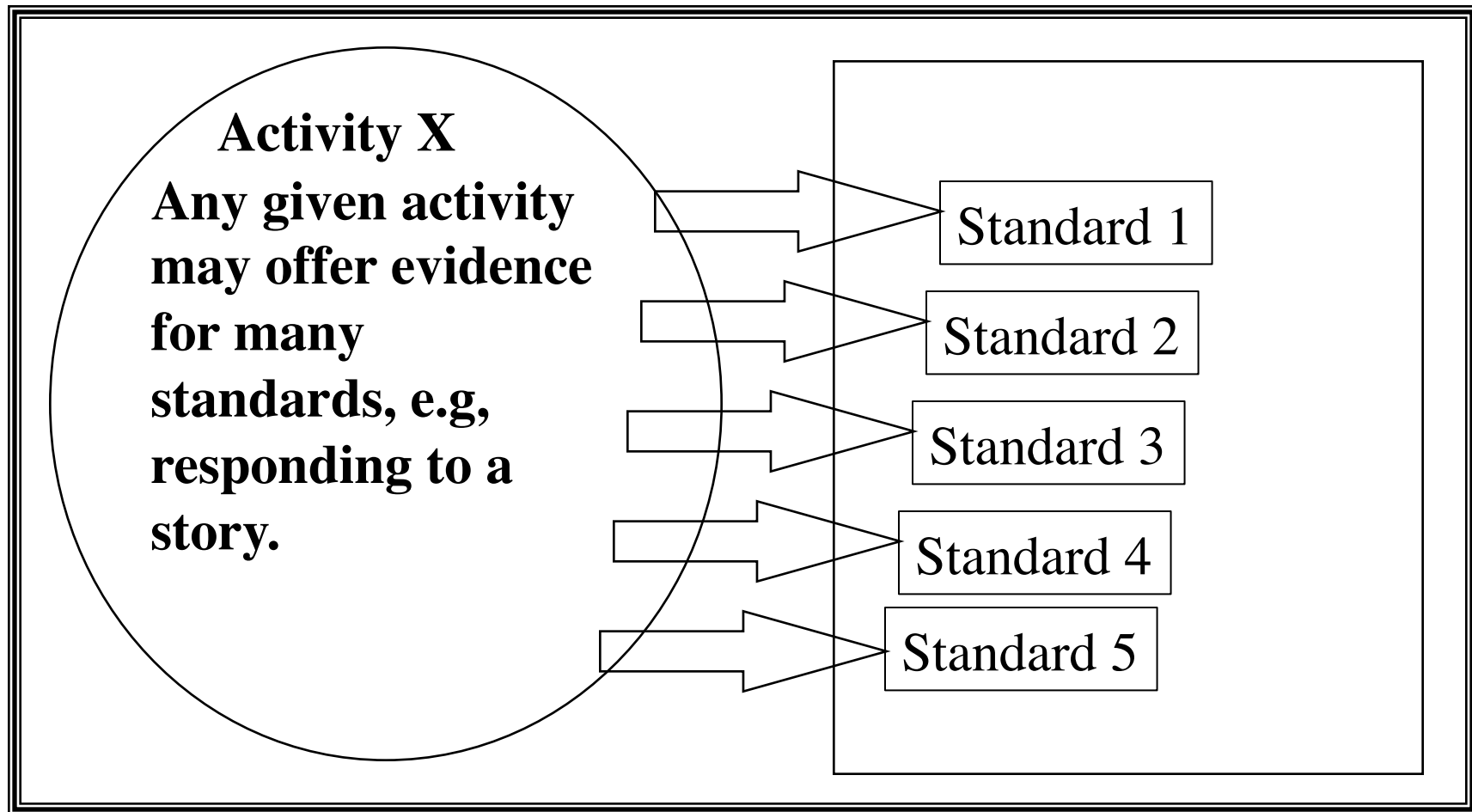
From

Conclusions

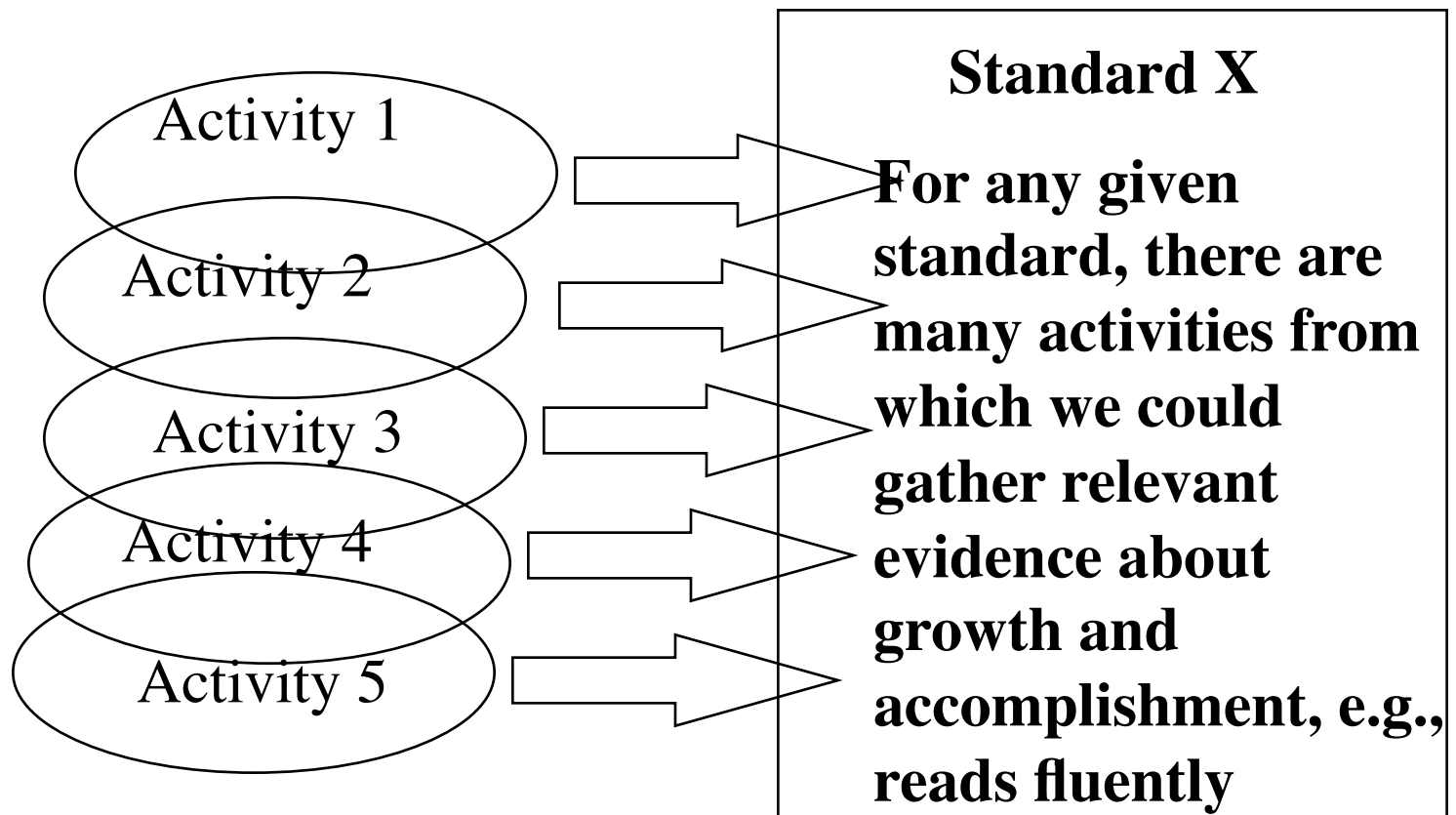
On standards 1-n

We engage in instructional activities, from which we collect evidence which permits us to draw conclusions about student growth or accomplishment on several dimensions (standards) of interest.

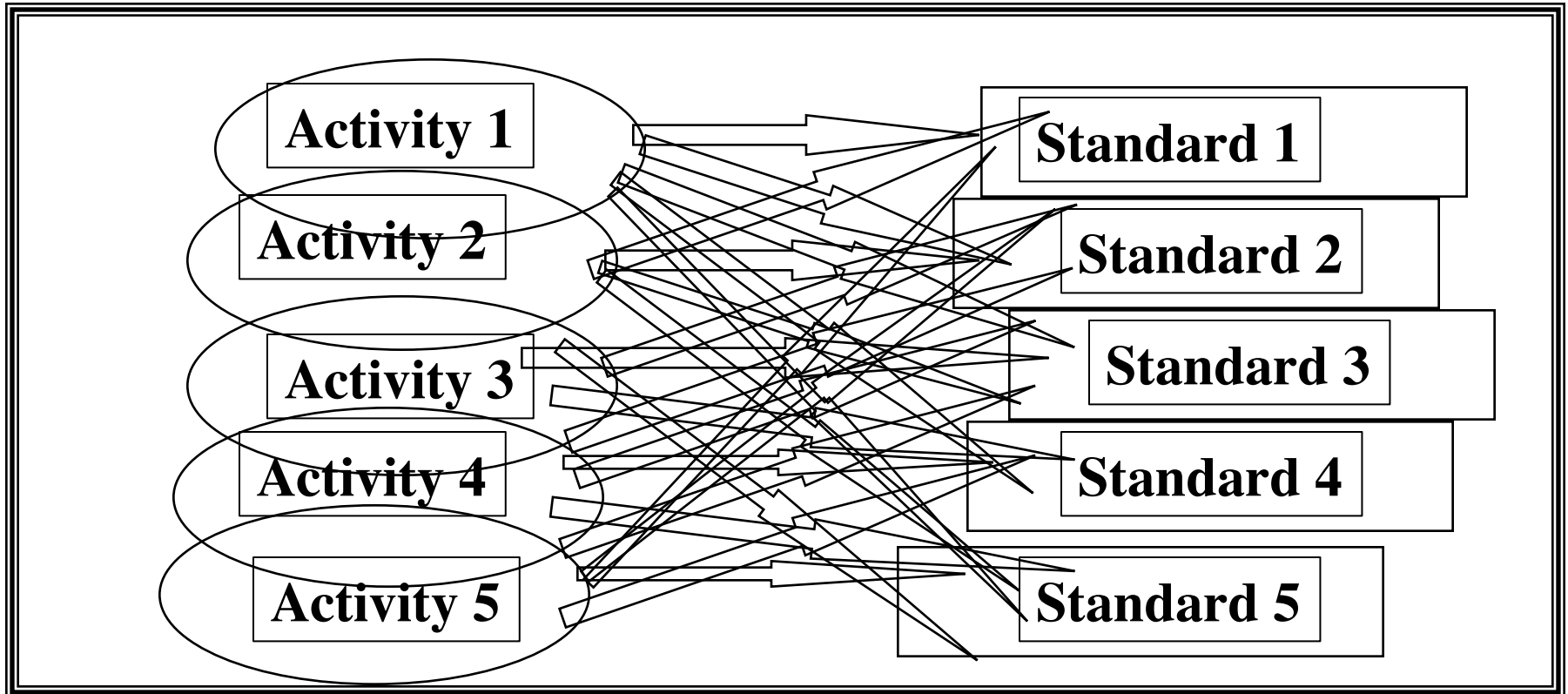
➤ The complexity of performance assessment practices: one too many



➤ The complexity of performance assessment practices: many to one



➤ The complexity of portfolio assessment practices, many to many



- artifact/activity can provide evidence for many standards
- Any given standard can be indexed by many different artifacts/activities

Example of a New Standards Task from mid 1990s

- Man and His Message
- MLK
- 6-8 days, depending on class time
- Culminating task: write an essay based upon choosing one of several prompt options.

Texts Encountered

- A video about the Civil Rights Movement entitled *Justice*.
- An article about the Civil Rights Movement entitled, *Confrontations*.
- An article about Ghandi from Scholastic's SEARCH magazine.
- An oral rendition of King's Letter from a Birmingham Jail.
- Printed versions of other King speeches.
- An excerpt from a Time magazine account of the Rodney King riots in East Central Los Angeles.
- Two CNN video accounts of the riots: *Rage of Despair* and *Roots of the Problem*.

Tasks Completed over the Period

- collaboratively complete separate cognitive webs on key concepts from the readings (Martin Luther King, Civil Rights Movement, Non-Violent Resistance).
- keep an ongoing log/chart of emerging learnings from all the different texts (written, oral, or video).
- answer straightforward "assignment-like" questions.
- compare the similarities and differences between King and Ghandi in a modified Venn diagram display.
- the culminating essay
- EVERYTHING can be scored

Affordances

- Has the look and feel of powerful or deeper learning
- Engages students in workplace like behaviors, including social behaviors
- Expands our conceptualization of what counts as a text
- High capacity for engagement: interest and relevance
- Maps onto many of the Common Core Standards for reading in History
- Could build professional community of teachers around implementation and scoring

The system imploded, exploded, reploded

- It died a premature death
- Three factors
 - Psychometric issues
 - Reliability
 - Task Generalizability
 - Measurement error
 - Cost issues
 - Teacher time
 - Professional Development
 - Political issues
 - Intrusion into personal space and values

The perils of performance assessment: or maybe those multiple-choice assessments aren't so bad after all.....

- Some people can tell what time it is by looking at the sun, but I never have been able to make out the numbers.
- There are four seasons: salt, pepper, mustard, and catsup.

The perils of performance assessment

- "Water is composed of two gins, Oxygin and Hydrogin. Oxygin is pure gin. Hydrogin is gin and water."

The perils of performance assessment

- "Germinate: To become a naturalized German."
- "Vacumm: A large, empty space where the pope lives."

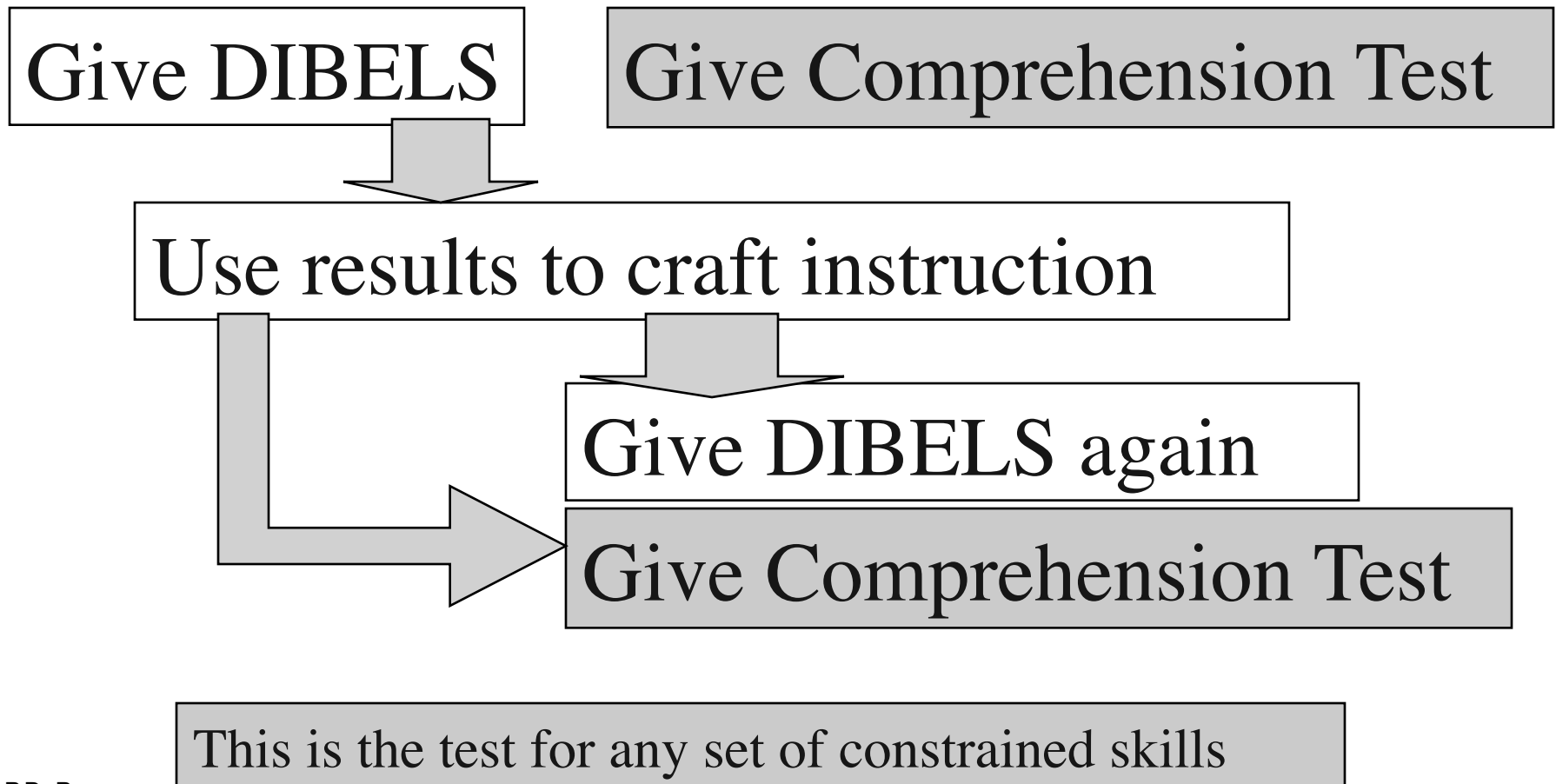
The perils of performance assessment

- Genetics explains why you look like your father, and if you don't, why you should.

The NCLB Years

- Retreat from performance assessment
- Measure and teach reading in its atomistic, not its global, aspect
- Link instruction to assessment by
 - Specifying all the specific standards
 - Building assessments for each standard
 - Making our instruction look just like the assessments
 - The DIBELS phenomenon

DIBELS Achilles Heel: Consequential Validity



My inference in 2009: Rule of Thumb for the Future

- Wherever we go, we must not go back to the world in which we link instruction to assessment at the very finest grain-size
- Keep our eyes on the prize—what does accomplished performance look like?
- My version: Return to the mid 90s and get it right this time...

www.scienceandliteracy.org

Where are we headed in the Post NCLB era?

- Obama administration keeping up the accountability pressure
 - Teacher accountability for student achievement
 - \$350,000,000 investment in new accountability assessments (PARCC, SBAC)
- CCSS movement
- Another round of curriculum reform
 - We are a part of that effort

A New Opportunity: A Confluence of Forces

- CCSS movements/Foundation funding
 - Disciplinary delivery of literacy
- Reincarnation of Ambitious Pedagogy
 - Deeper Learning
- PARCC and SBAC
 - Hybrid assessments
 - Increase weights given to open ended responses and complex performance tasks

Why performance assessment?

Performance Tasks: Why bother?

- External validity
 - College ready
 - Career ready
 - PARCC and SBAC ready
- Curricular validity
 - Higher Order Thinking
 - Powerful learning
 - Deeper learning

Performance Tasks: Why bother?

- Consequential validity
 - What curricular activities will it lead teachers and students toward?
 - Scaffolding toward independence
- Utility
 - Formative Assessment
 - Feedback to the student→differentiated scaffolding
 - Feedback to you as the teacher→reteaching
 - Summative
 - Did we achieve our goals?
 - Interim—
 - Are we meeting benchmarks along the way

Shifts in Emphasis

PARCC: What's new?

- **Complexity:** The standards require regular practice with **complex text** and its academic language
- **Evidence:** The standards emphasize reading and **writing grounded in evidence** from text, both literary and informational
- **Knowledge:** The standards require **building knowledge** through content rich non-fiction

Smarter Balanced: What's new?

- Complex way of mapping items to the standards
- Claims & Targets
- Lots of Rubrics, Anchor papers
- Complex Tasks that provide scores for reading and writing
- Formative Assessments Galore

SBAC Claims are VERY Global

- Claim #1 – Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts.
- Claim #2 – Students can produce effective and well-grounded writing for a range of purposes and audiences.
- Claim #3 – Students can employ effective speaking and listening skills for a range of purposes and audiences.
- Claim #4 – Students can engage in research/inquiry to investigate topics, and to analyze, integrate, and present information.

Common Elements

- Hybrid Formats
 - Selected-response (SR) items,
 - constructed-response (CR) items,
 - performance tasks (PT), and
 - technology enhanced (TE) items/tasks
- Design process
 - Evidence Centered Design
 - Elaborate Process based in cognitive science and curricular learning progressions

PARCC exemplars...

- Karen's presentation this afternoon will feature lots more examples plus the logic of the assessments

Grade 6 EBSR from Narrative Writing Task

SAMPLE ITEM

Part A

Based on the passage from *Julie of the Wolves*, how does Miyax feel about her father?

- ☐ a. She is angry that he left her alone.
- ☐ b. She blames him for her difficult childhood.
- ☐ c. She appreciates his thorough knowledge of nature.
- ☐ d. She is grateful that he planned out her future.

Part B

Which sentence from the passage best shows Miyax's feelings for her father?"

- ☐ a. "She had been lost without food for many sleeps on the North Slope of Alaska."
- ☐ b. "This could be done she knew, for her father, an Eskimo hunter, had done so."
- ☐ c. "Unfortunately, Miyax's father never explained to her how he had told the wolf of his needs."
- ☐ d. "And not long afterward he paddled his kayak into the Bering Sea to hunt for seal, and he never returned."

Passage

George, Jean C. *Julie of the Wolves*. New York: Harper and Row, 1972. Print.

PARCC is committed to using authentic texts. Permissions are pending for the texts associated

Grade 6 TECR from Narrative Writing Task

SAMPLE ITEM

Part A

Choose one word that describes Miyax based on evidence from the text. There is more than one correct choice listed below.

- ☐ reckless
- ☐ lively
- ☐ imaginative
- ☐ observant
- ☐ impatient
- ☐ confident

Part B

Find a sentence in the passage with details that support your response to Part A. Click on that sentence and drag and drop it into the box below.

Part C

Find another sentence in the passage with details that support your response to Part A. Click on that sentence and drag and drop it into the box below.

Passage


George, Jean C. *Julie of the Wolves*. New York: Harper and Row, 1972. Print.

SAMPLE ITEM

In the passage, the author developed a strong character named Miyax. Think about Miyax and the details the author used to create that character. The passage ends with Miyax waiting for the black wolf to look at her.

Write an original story to continue where the passage ended. In your story, be sure to use what you have learned about the character Miyax as you tell what happens to her next.

Answer:

Font Size... ▼	Font Family. ▼	B	<i>I</i>	<u>U</u>	≡	≡	≡	≡	≡	≡	X ₂	X ²	

Research Project

- Each module includes the opportunity for students to produce one **extended project** that uses research to address a significant topic, problem, or issue.
- This entails gathering and synthesizing relevant information from several additional literary or informational texts in various media or formats on a particular topic or question drawn from one or more texts from the module.
- Students are expected at this stage to assess the credibility of each source, effectively and accurately quote or paraphrase sources, and include basic bibliographic information in their research.
- Students can present their findings in a variety of informal and more formal argumentative or explanatory contexts, either in writing or orally. (Research aligned with the standards could take one to two weeks of instruction.)
- Ongoing incorporation of research for shorter tasks should also be a regular component of instruction.

Smarter Balanced Exemplars

- Sample Constructed Response Item
- Claim 1, Assessment Target 11: *Item Prompt: Write a paragraph explaining why people who live in moist climates work harder to prevent mold than people that live in dry climates. Include details from the passage about how they prevent mold.*

Performance Task Overview (120 total minutes): Title: Virtual Schools

Part 1 (35 minutes): Ultimately tasked with writing an argumentative essay on virtual schools, students will read a story and article, view a video, and review research statistics, taking notes on these sources. They will then respond to three constructed-response questions addressing the research skills of analyzing and evaluating information.

Part 2 (85 minutes): Finally, students will work individually to compose full-length argumentative essays on virtual schools, referring to their notes as needed. Pre-writing, drafting, and revising will be involved. Scorable Products: Student responses to the constructed-response questions and the essay will be scored.

Technology enhanced

- **Select three sentences that show that Naomi is worried she has done something wrong.**

Using text details to support an inference or generalization...

From Smarter Balanced.

Why There Should Be a Longer School Day

Schools should have a longer school day for students. First, students could learn more about different subjects if the school day were longer. Also, students could get extra help from teachers. More hours in class each day would also mean more vacations scattered throughout the year!

- Now look at the following daily schedule for a school that has switched to a longer school day.

- 8:00 Morning Announcements
- 8:20 Reading Language Arts
- 9:30 Foreign Language
- 10:30 Morning Recess
- 10:45 Mathematics
- 11:45 Lunch
- 12:45 History
- 1:45 Art or Music
- 2:15 Afternoon Recess
- 2:45 Science
- 3:30 Homework Preparation
- 3:45 After-School Tutoring or Sports
-

Even the writing prompts entail close reading as careful examination of the text

- Revise the paragraph by adding details from the daily schedule that help support the reasons for having a longer school day.

My analysis

- Students who have learned how to read and write in curriculum that requires constructed responses and real writing will lead to good performance on PARRC and SBAC assessments
- Students still wallowing in a bits and pieces curriculum will not.

Why?

- They will have developed some transferrable practices that will serve them well in these new circumstances.
- That's what we are about—developing transferable knowledge and skills

Some good work...

- New Visions Schools in NYC, along with LDC and SCALE at Stanford
- Lucy Calkins Reading-Writing Workshop
- Common Elements
 - True to CCSS
 - True to the Vision of Project Based Learning
 - Challenging curriculum and pedagogy
- Literacy Research Panel initiative...

Teaching Task: Culminating Summative Assessment

Industrial Revolution Module



23 texts over 3 weeks

Were the achievements of the Industrial Revolution worth the costs to society?

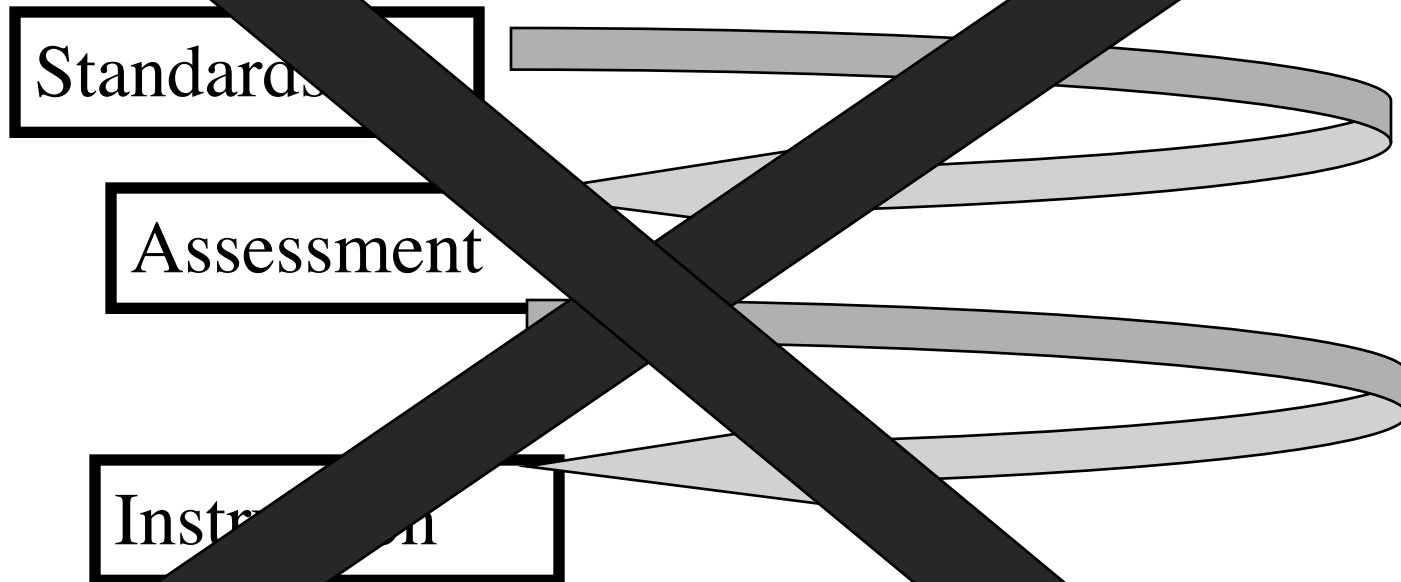
Lots of practice of specific standards
Writing from reading sources

After reading secondary and primary sources pertaining to the British Industrial Revolution, write an argumentative essay that addresses the question and support your position with evidence from the texts. Be sure to acknowledge competing views.

Closing Advice...

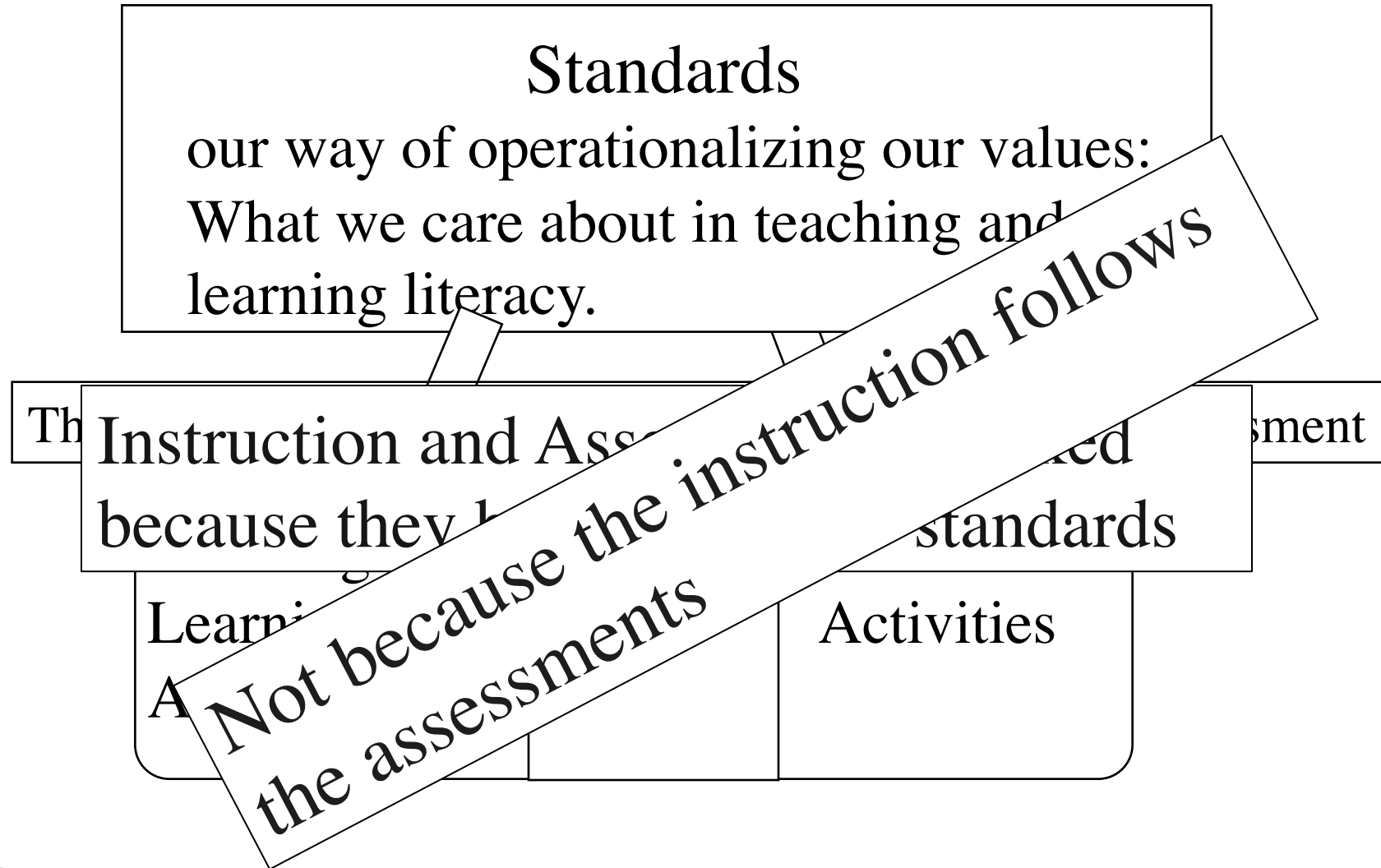
- Be careful about how we link instruction and assessment

A model worth rejecting: Aligning assessment to standards and instruction to assessment



- This model is likely to shape the instruction narrowly.
- Leads to students who can pass the tests but not perform the cognitive tasks the tests are supposed to measure.

A better way of thinking about how standards can link instruction and assessment



High Hopes!!!!

- On good days, I think that CCSS. Filtered through PARRC and SBAC, are up to the challenge
- They are certainly light years ahead of the current generation of state standards and assessments
- I'll be watching with cautious optimism...

My bottom line

- We desperately need *instructionally sensitive* assessments that have first rate *psychometric* characteristics so that we can build *trustworthy* internal systems for monitoring student progress
- No decision of consequence about any individual, school, district or other aggregation should be based upon a single indicator of anything.

Bottom Line

- Tests are a means to an end: Their value is indexed by the degree to which they allow us to make good decisions about **and** provide good instruction for kids, classes, and schools.
- Bad things happen when tests become ends themselves.

Bottom Line...

- Stakes matter!
- Challenge matters!
- High stakes and low challenge is the worst possible scenario...
 - Least common denominator
 - Poor kids will be stuck in a cycle of basic skills

The real bottom line

- Never send a test out to do a curriculum's job!

Instead, send the test out to support
the curriculum and those held
accountable to it

