

# Teaching and Testing in the Common Core Era

## *The Shifting Testing Landscape*

Education Writers Association Seminar  
*Los Angeles, CA*  
February 26, 2016

***Chris Domaleski***  
*The National Center for the Improvement of  
Educational Assessment*

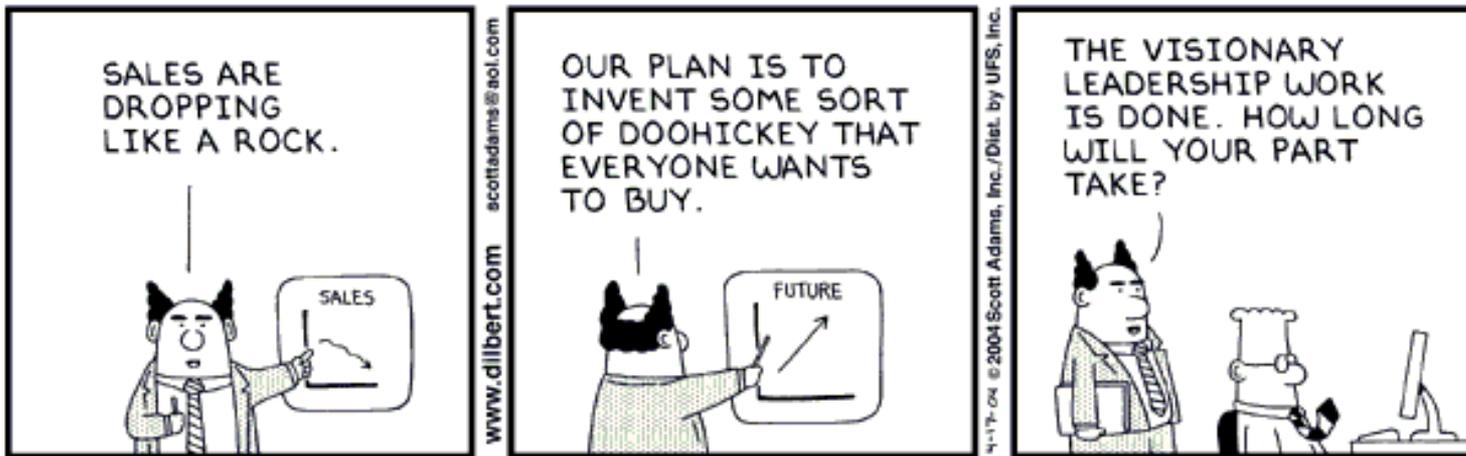
# The Shifting Landscape

- My focus today:
  - Discuss stakeholder priorities for assessment and how this is influencing the 'shifting landscape'
  - Identify some implications of competing priorities to consider
  - Suggest some key questions moving forward

# Assessments: what do stakeholders want?

- Useful feedback to inform instruction
- Can be used to certify proficiency with respect to state standards and used for various accountability purposes
- Measure cognitively rich and relevant knowledge and skills
- Compare performance at student and aggregate levels
- Timely (preferably, immediate) results
- Low cost
- Minimal testing time and operational burden
- Ready now

# This is easy, right?



© UFS, Inc.

# Competing Priorities

- Obviously, the path forward will involve finding the balance point among competing priorities
- Let's unpack a few...

# Autonomy and Comparability

- Many states feel strongly about protecting the autonomy to define what will be measured and how. For example:
  - Recent and ongoing changes to state content standards
  - Desire to exercise choice in operational providers and decisions
- Many states continue to advocate for a focus on comparability. For example:
  - Cross-state comparability of student or aggregate results
  - Reference to key benchmarks such as NAEP, ACT, SAT, and/or international tests

# Improved Information and Efficiency

- Stakeholders increasingly want more information from assessments. For example:
  - Measure deeper and broader range of skills associated with post-secondary success
  - Reports are useful to inform instruction
- There is increasing push-back related to test burden and cost. For example:
  - Loss of instructional time due to frequency and/or length of assessments
  - Swelling costs primarily associated with development and scoring constructed response

# Important Questions

- What is the timeline and process for transition? For example:
  - Are items acquired from extant sources, custom developed, or both? If the former, what is the process for ensuring the items are aligned and appropriate? When and how are items reviewed and piloted?
  - When and how are operational systems (e.g. computer based testing platforms) evaluated and certified?
- What evidence will be available to support assessment claims? For example:
  - United States Department of Education Peer Review
  - Independent evaluations (e.g. TAC review)
- What comparability claims are being made and how are these claims supported?
  - Are current year results comparable to prior year results? Do comparability claims differ for status and growth?
  - Are results within state comparable to results from other states or programs? It is important to consider that comparability is based on much more than similar blueprints (e.g. administration and scoring have substantial impact).
  - What steps are taken to establish comparability between administration modes (e.g. CBT to PBT, CBT across devices)?

# Promising Practices

- Focus on coherent assessment systems to support multiple purposes in lieu of ‘overloading’ any one test
  - *If we ask a test to be useful for everything, it likely won’t be useful for anything*
- States exploring new models for sharing content and collaborating
  - For example, recently announced ‘tier structure’ of PARCC
  - States working together with common providers outside consortia

# Contact Information

- Chris Domaleski
  - Associate Director, Center for Assessment
  - [cdomaleski@nceia.org](mailto:cdomaleski@nceia.org)