Center for Assessment

ccee
California Collaborative
for Educational Excellence

A Tricky Balance: Introduction to Balanced Systems of Assessment

Scott Marion and Carla Evans

National Center for the Improvement of Educational Assessment

Webinar #1 of the Webinar Series: *Developing and Implementing Balanced Assessment Systems to Support School Improvement and Student Learning*

California Collaborative for Educational Excellence

April 1, 2021

Introduction (This session is being recorded)

- This is the first of **five webinars** offered in conjunction with **CCEE** designed to support California district leaders and others in designing, developing, and implementing balanced assessment systems to support multiple purposes and uses.
- **Today**:
  - ➜ What's the problem?
  - ➜ Why balanced assessment systems?
  - ➜ How do we tell if it is balanced?
  - ➜ What's getting in the way?
  - ➜ What can we do about it?
- We're going to use the chat and some polls to keep you on your toes!

National Center for the Improvement of Educational Assessment

- A non-profit technical & policy consulting firm established in 1998 with the mission of improving student learning through improved assessment and accountability practices
- Current contracts with 40+ states, districts, and other entities
  - Almost all are long-term contracts designed to provide technical and design support for a range of assessment and accountability issues
- Purposely small—13 full-time professionals
  - All with doctoral degrees and almost all have worked in the “real world” of state assessment and/or as assessment contractor staff.
- Non-partisan and independent of any governmental agency or testing company
- Recognized as leaders in the design and implementation of local and state assessment systems

Why Balanced Assessment Systems?

What "problem" are we trying to solve?

Why do you think—based on what you know now—that having a balanced (or more balanced) assessment system will address?

If we are not clear about what problem/issue we're trying to solve/address, no design will help us.

In the chat to <u>everyone</u>, please jot down some of the issues or problems you'd like to address in your current assessment system…

Assessments to Support Instruction

- From prior experience, we expect many of you to say “we want assessments to help teachers and students improve teaching and learning”
- What assessment characteristics are necessary for it to support improvements in instruction and learning?
- **Again, post some design characteristics in the chat**. Please be as specific as possible…

5

My Five Critical Characteristics*

1. Coherent with the enacted curriculum
   - Standards are not fine-grained enough
2. Items and tasks must foster and support deeper thinking
3. Results at the right grain size to support useful feedback
4. Timely results
   - Not necessarily instant, but soon enough to adjust instruction when it matters
5. Results to inform instruction—helps inform what to do next
   - This is much harder to do than it sounds

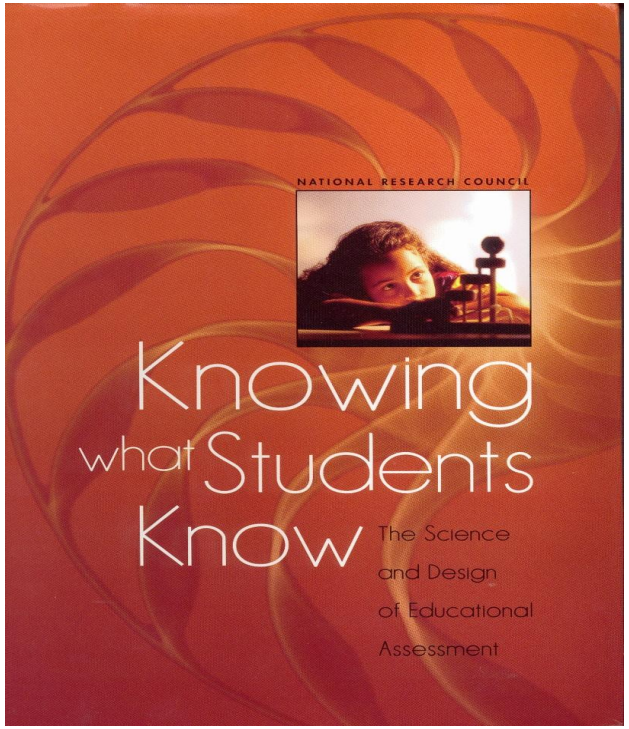
***See: Five Essential Features of Assessment for Learning**

Assessment Design Involves Tradeoffs

A key trade-off in current assessment design: Accountability versus instructional support and improvement for individual students

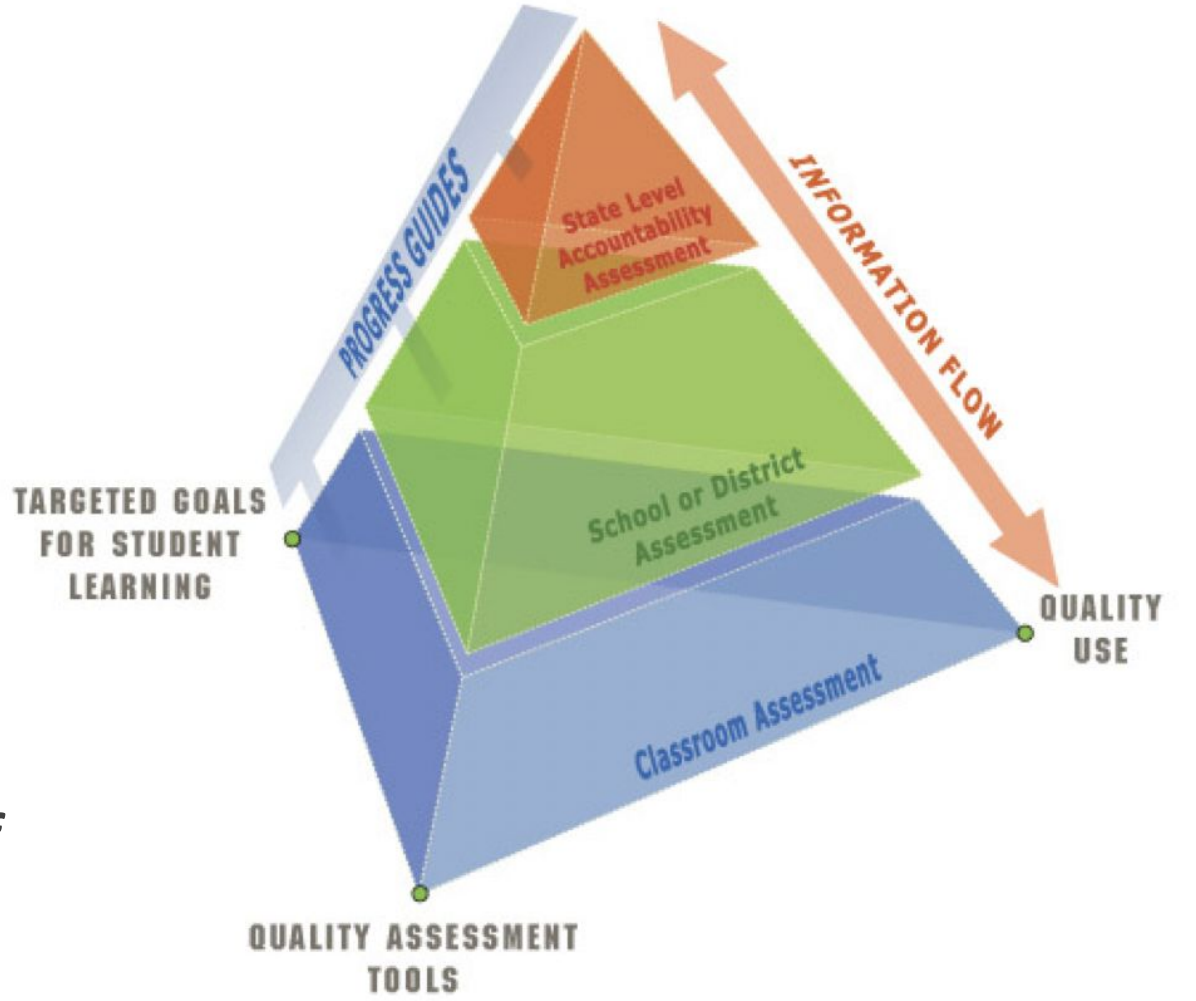
"Ironically, the questions that are of most use to the state officer are of the least use to the teacher." (National Research Council, 2001)

Why? Timing, grain size, connection to the enacted curriculum…



A Call for Balanced Assessment Systems

*Assessments at all levels—from classroom to state—**will work together in a system that is comprehensive, coherent, and continuous**. In such a system, assessments would provide a variety of evidence to support educational decision making. **Assessment at all levels would be linked back to the same underlying model of student learning** and would provide indications of student growth over time* (NRC, 2001, p.9).



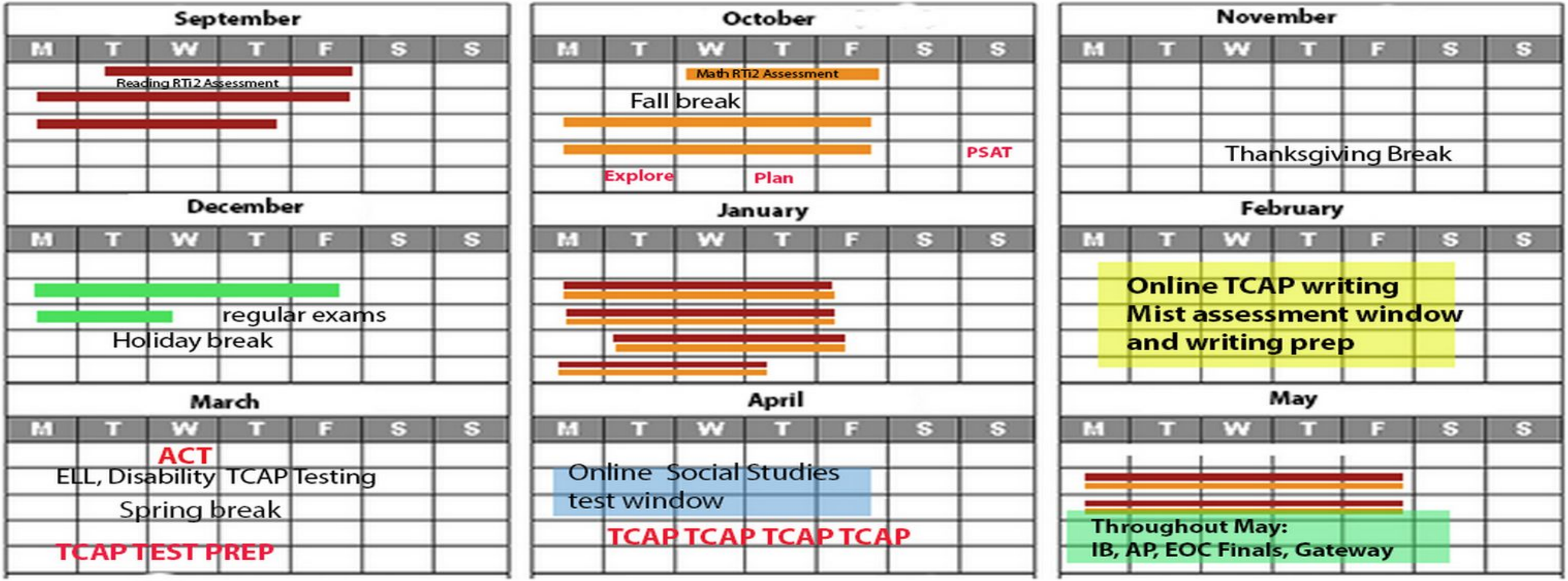
From Pellegrino, 2019

…a collection of assessments does not entail a system any more than a pile of bricks entails a house.

Coladarci, T. (2002). Is it a house...or a pile of bricks? Important features of a local assessment system. *The Phi Delta Kappan*, 83(10), 772–774. Retrieved from http://www.jstor.org/stable/20440251

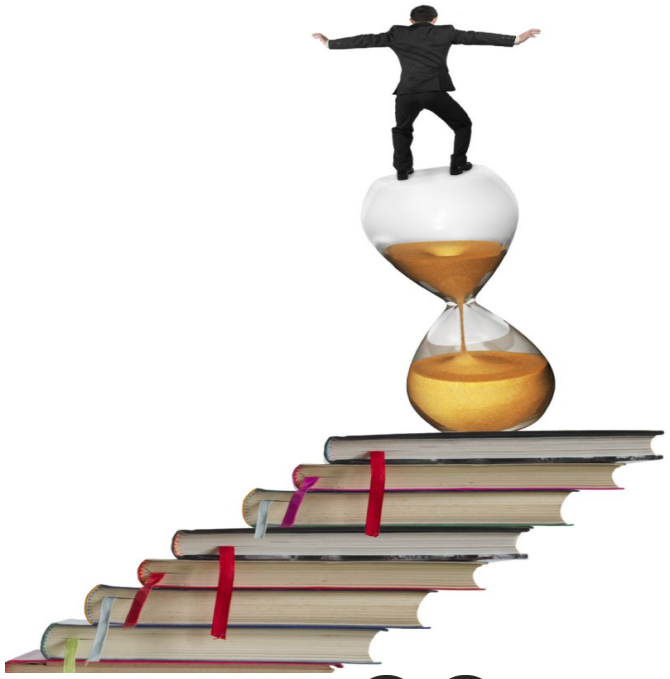
This is a real assessment calendar!

In the **poll**, please answer if your assessment calendar looks **better**, **worse** or about the **same** as this slide?



Balanced Assessment Systems to Serve Multiple Purposes

Assessment systems designed to serve multiple purposes require thoughtful planning about which data will be privileged at each level (Chattergoon & Marion, 2016). For example:



- standardized vs. dynamic/flexible
- uniform vs. variable dates
- independent vs. assisted (scaffolded)performance
- delayed vs. immediate feedback
- stringent requirements for technical accuracy vs. less stringent requirements

Quality assessments are needed at all levels!

What we *can* assess on larger scale district or state assessments

What we *should* assess on classroom and school assessments

Criteria for Evaluating Balanced Assessment Systems

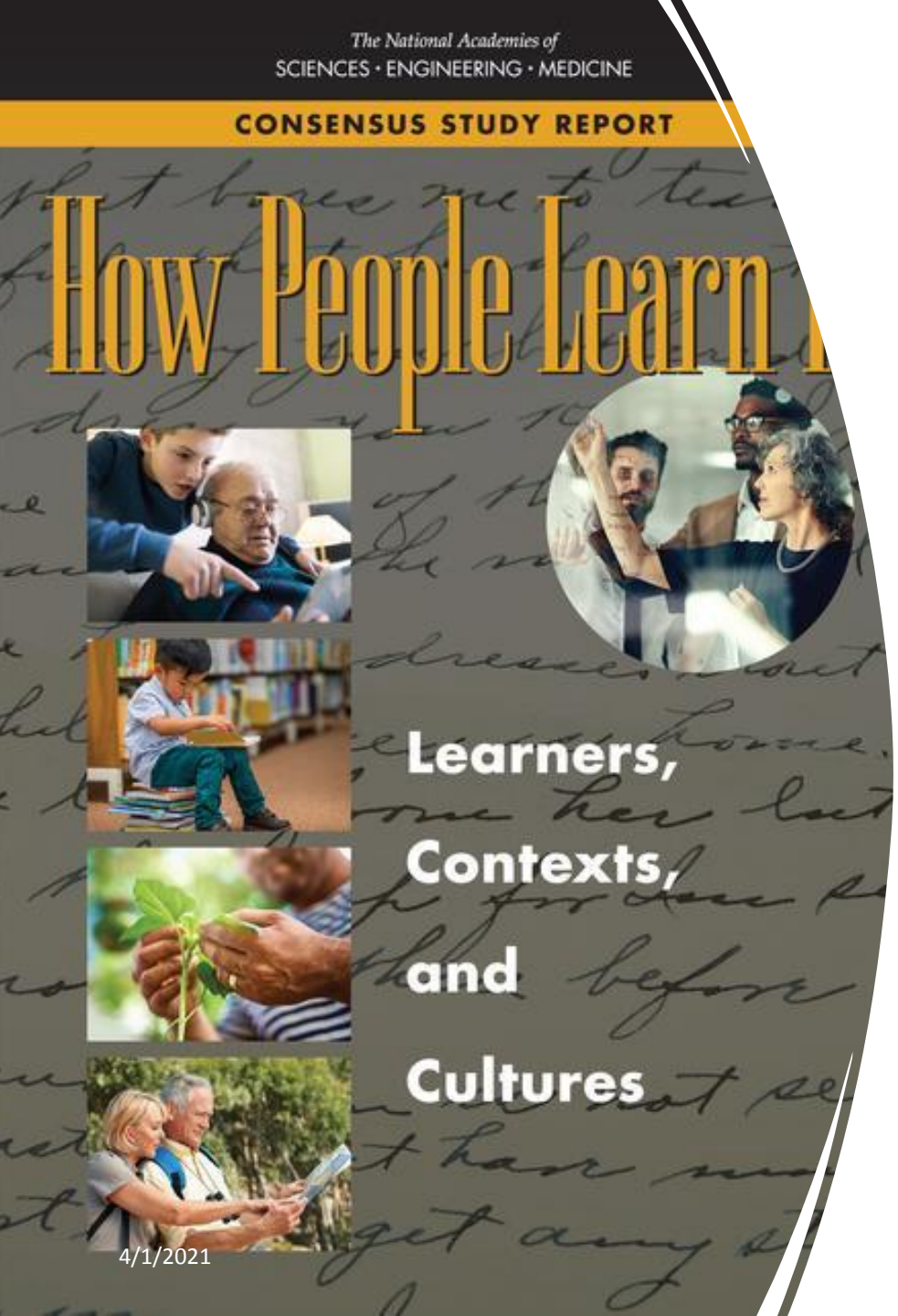
- A **balanced** assessment environment should exhibit three properties (NRC, 2001):
  1. **Comprehensiveness** – "a range of measurement approaches should be used to provide a variety of evidence to support educational decision-making"
  2. **Coherence** – "the conceptual base or models of student learning underlying the various external classroom assessments within a system should be compatible"
  3. **Continuity** – "assessments should measure student progress over time"

Two additional criteria

4. **Efficiency** means getting the most out of assessment resources and eliminating redundant, unused, and untimely assessments.
5. **Utility** is a key criterion for assessment system quality should be the degree to which the system provides the information necessary to support the intended aims
    - Follows from a well-articulated theory of action that specifies the various intended outcomes for the system and the processes and mechanisms by which these outcomes will be realized.

Coherence

- Vertical **Coherence** – conceptual base or models of student learning underlying the various **external** and **classroom** assessments within a system should be compatible
- Horizontal **Coherence** – alignment among curriculum, instruction, and assessment along a common set of learning goals



Coherence - Not Just Any Model of Learning

Assessments and assessment systems must be based on research-based models of learning.

Adherence to outdated, naïve, and/or implicit notions of learning are an impediment to assessment literacy and assessment reform.

A huge PD issue!

National Academies of Sciences, Engineering, and Medicine. 2018. *How People Learn II: Learners, Contexts, and Cultures*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24783.

It's All About Utility!

u·til·i·ty
/yo͞o ˈtiləd ē/

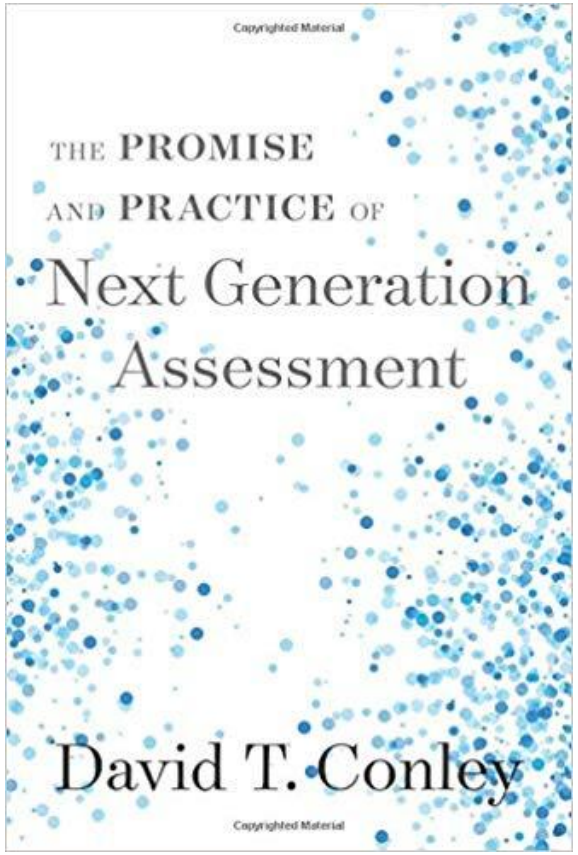
noun

1. the state of being useful, profitable, or beneficial.
   "he had a poor opinion of the utility of book learning"
   synonyms: usefulness, use, benefit, value, advantage, advantageousness, help, helpfulness, effectiveness, avail; *formal* efficacy
   "we have increased the machine's utility"
2. a public utility.

adjective

1. useful, especially through being able to perform several functions.
   "a utility truck"
2. functional rather than attractive.
   "utility clothing"

Real or Mythical?



- We've been talking and writing about balanced assessment systems for over 20 years
- Are balanced assessment systems the unicorns of educational assessment?
- David Conley went searching for examples of balanced assessment systems for his new book, but found only partial examples
- Why?

Barriers to Balance

Politics and Policy

Weak ties to curriculum & learning

Assessment literacy

Weak Design Thinking

Proliferation and commercialization

Many more, but we think these are the “big 5”

Weak ties to curriculum & learning

- *Knowing What Students Know* and many documents that followed stressed that a shared model of learning was critical for vertical and horizontal coherence.
- The state test aligns to common content standards; aren't these standards sufficient for designing a balanced assessment system?
- Why has this been a barrier to the development of balanced assessment systems?
- What level of specificity in the learning model is necessary to support vertical and horizontal coherence?
- Relatedly, most states are curriculum agnostic. How has this approach to curriculum affected the development of coherent systems and what can be done about it?

Webinar #2: The critical role of curriculum and learning progressions in balancing assessment systems - **Tuesday, April 13th (3-4pm PST)**

The second webinar focuses on the most important criterion—**coherence**—and how it can be supported with high-quality curriculum and learning progressions. Relying on experts in curriculum, this webinar highlights the features that distinguish high-quality curriculum and learning progressions from lower quality processes and products.

Participating experts:

- **Rebecca Kockler**, Education First and former Louisiana Deputy Commissioner for Teaching and Learning
- **Jeri Thompson**, Senior Associate, Center for Assessment
- **Scott Marion**, Executive Director, Center for Assessment

Click here to register in advance for this webinar.

Rebecca Kockler

Jeri Thompson

Proliferation & commercialization

- Since NCLB was passed in 2001, state and district assessments, especially interim assessments, have proliferated.
- Let's ask the pink elephant question:
- What would commercial interim assessments have to look like (and do) to play a useful role in a balanced assessment systems?
- Do commercial interim assessments have a role in balanced assessment systems?

ccee
California Collaborative
for Educational Excellence

Center for Assessment

Webinar #3: The components of a balanced assessment system

Tuesday, April 20th (3-4pm PST)

There is an unfortunate "urban legend" about balanced assessment systems. Many appear to think that an assessment system is balanced when it includes summative, interim, and formative components. This is not true. This webinar discusses the key features of various types of assessments and helps participants understand how to evaluate components of an assessment system considering the intended purposes and uses. We discuss, specifically, what it takes for assessments to be instructionally useful as well as the power of curriculum-embedded, unit-based assessments.

With **Lorrie Shepard**, Distinguished Professor, University of Colorado Boulder

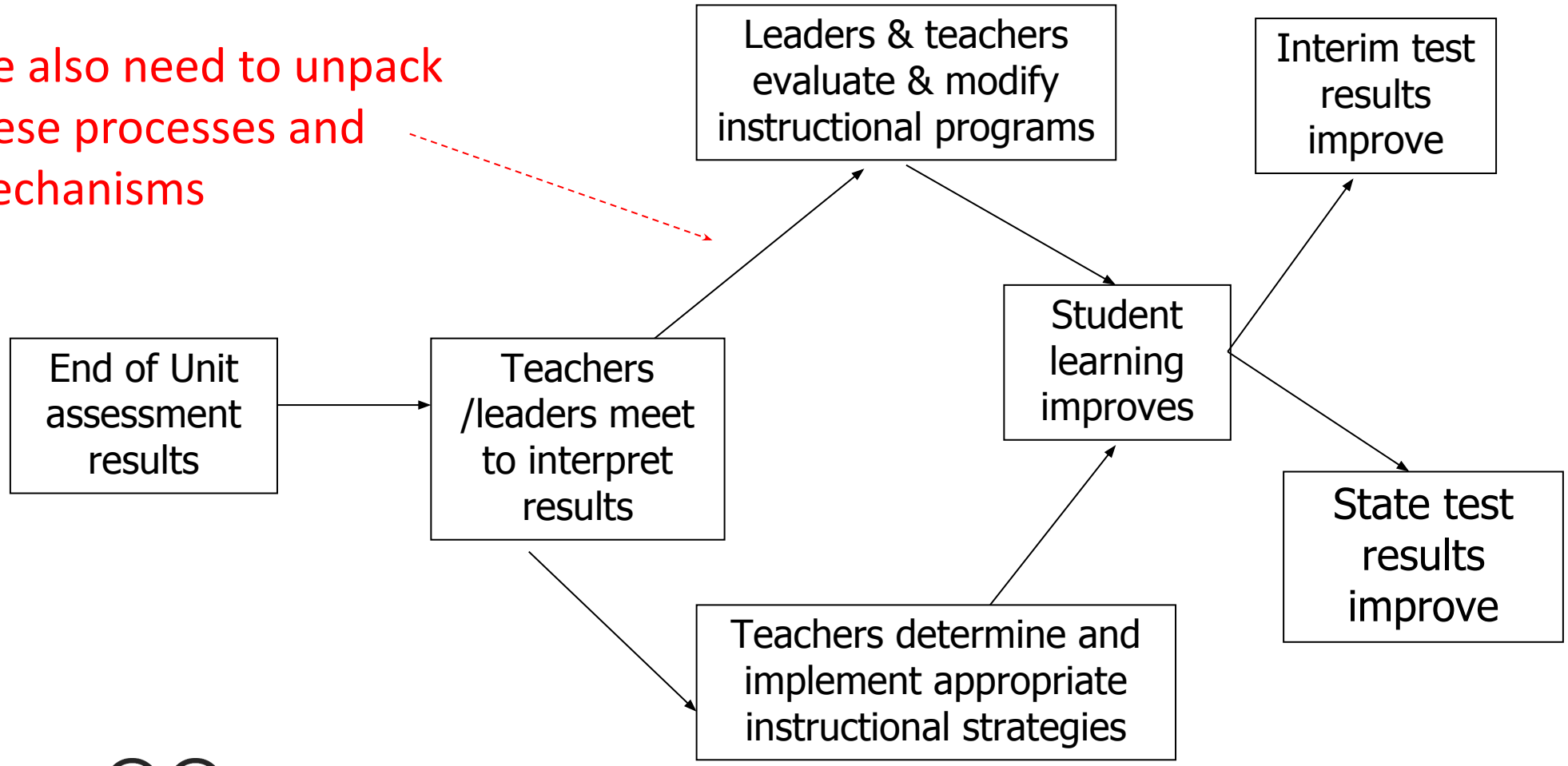
Click here to register in advance for this webinar.

Weak Design Thinking

- We and others have long been concerned that many people engage in design activities without thinking through how all the pieces fit together.
- This is a concern for single assessments that is heightened when thinking about how to turn a pile of bricks into a house or how to put Humpty Dumpty together again
- Theories of action or logic models have been very helpful to guide design, implementation and evaluation.

A Simple Theory of Action Hypothesizing How the Use of End-of-Unit Assessments Lead to Improvements in State Test Scores

We also need to unpack these processes and mechanisms



Webinar #4: Theories of action as a tool for developing and implementing balanced assessment systems—**May 11th**

In webinar #4, we introduce theories of action as a tool for helping district and school leaders identify the problems/issues they are trying to address through the design and implementation of assessment systems. Further, theories of action provide a comprehensive framework for analysis, evaluation, and continuous improvement. In addition to helping participants understand the usefulness of theories of action, we share several practical tools to help participants begin to develop theories of action for balanced systems of assessment.

- Participating experts:
  - Nathan Dadey, Senior Associate, Center for Assessment
  - Erika Landl, Senior Associate, Center for Assessment

•Register in advance for this webinar:

https://us02web.zoom.us/webinar/register/WN_wzdYDw6oR-WtLsIOPNm5pw

Politics and Policy

- Who controls what?
  - States, districts, and schools "own" or control different assessments so how do we turn a pile of bricks into a house?
  - Essentially all states are curriculum agnostic, so can we have a coherent state assessment system?
- Two main policy considerations:
  - What is the influence of state accountability policies on designing and implementing state assessment systems?
  - How have the policies associated with the use, instability (often), and footprint of state assessments affected the implementation of balanced assessment systems at the state and/or district level?

Webinar #5: Does it quack like a duck? Would I know a balanced assessment system if I saw one? May 27th

Thursday, May 27th (3-4pm PST)

This culminating webinar will include examples and partial examples from the field to help participants understand what developing systems look like in practice. Participants will hear from district assessment leaders across the country from various size school districts to understand how they are adapting the theory and practice of balanced assessment system development to their contexts. Participants will gain a good understanding of the challenges faced by these various leaders and hear how they are trying to overcome them.

Participating experts:

- Peter Leonard, Director of Assessment, Chicago Public Schools
- Scott Marion, Executive Director, Center for Assessment
- TBA

Click here to register in advance for this webinar.

Assessment literacy

- Folks like Rick Stiggins have been talking about assessment literacy for more than 40 years.
  - In fact, there's an assessment literacy session going on now (don't leave!).
- What aspects of assessment literacy are unique, or at least particularly important, to the design and implementation of balanced assessment systems? How do these knowledge and skills differ for various stakeholders:
  - Teachers?
  - School leaders?
  - District leaders?
  - State leaders?

Barriers to Levers (some examples)

- Curriculum coherence
  - Create model curriculum, replacement units, or learning progressions
- Proliferation and commercialization
  - Support assessment quality evaluations; assessment audits/mapping activities; “loose-coupling”
- Design Thinking
  - Learn how to create theories of action and use other design approaches to guide design, implementation and evaluation
- Policy and politics
  - Reducing accountability pressure; avoiding narrow focus
  - Assessment policy to support stability and reducing the footprint of the state assessment
- Assessment literacy
  - Beyond the incredible challenge of addressing assessment literacy, states can support **assessment system literacy** to help district and school leaders understand what it takes to create a coherent system

Barriers to Levers (poll)

- Please **respond to the poll** and identify what you see as the primary barrier hindering the implementation of balanced assessment systems in your local entity.
- As you finish the poll, think about **what you can really do** to address this barrier and turn it into a lever.
- Please use the chat function to share your ideas.

Questions and responses…

What additional questions or comments do you have?

Contact us at:

Scott Marion: smarion@nciea.org or via Twitter @ScottFMarion
Carla Evans: cevans@nciea.org or via Twitter @CarlaMEvans

www.nciea.org
@NCIEA1

Extra slides

Purposes and Uses

- Each assessment is designed to serve a very limited number of purposes and uses (generally only one)
- Unfortunately, vague descriptions are not that useful for guiding assessment design or selection.
- Here are some examples of somewhat more specific purposes/uses:
  - My assessment will be used by district leaders to evaluate curricular and instructional programs.
  - My assessment will be used by district curriculum leaders to provide information about student learning gains when exposed to new curricular and instructional programs.
  - Teachers will use student work analyses protocols on student responses to performance assessments to understand how students are reasoning with new math concepts.

Comprehensiveness also addresses multiple users

Purposes/Uses

- Accountability
- Monitoring Equity
- Instruction/learning
- Grading
- Program/curricular evaluation

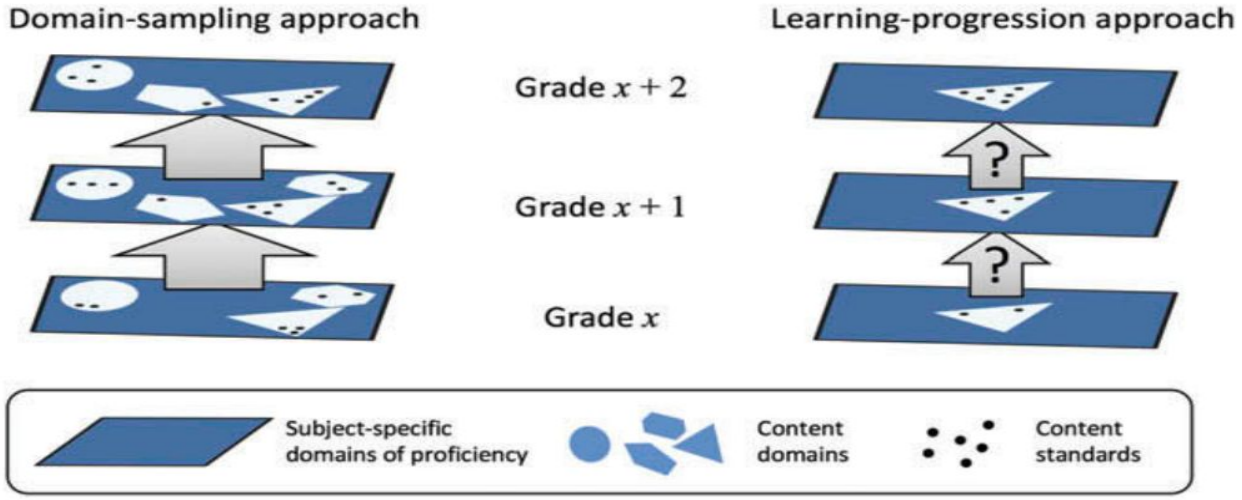
Context and users

- State policy leaders
- District leaders
- District CIA leaders
- Principals
- Teachers
- Students
- Parents

Assessments must be designed to support well-defined **purposes** and intended **uses**.

Continuity

- Continuity criterion can be met when progress is evaluated using same set of assessments (e.g., state tests)
- More challenging when trying to evaluate progress across assessments controlled by different entities and used for different purposes



From Briggs & Peck, 2015

A Call to Action

We sense a desperate need to improve the quality and usefulness of assessments. Balanced assessment systems have been proposed for meeting many needs, but we do not see enough examples of such systems in practice to serve as models for others to emulate. We named several key challenges that explain why such assessment systems are rare, and we suggested approaches for ameliorating some of these challenges. We would like to look back after the next 20 years and see more progress than we have seen in the almost 20 years since the publication of Knowing What Students Know.

What do we need to learn?

What do we need to learn and do to ensure we see coherent and cohesive systems that span from the classroom to the state?

- Conceptual challenges?
  - Criteria?
  - Purposes and uses?
  - Systems within systems
  - Components of the system (levels, layers, formats)?
- Practical?
  - Tools and resources
  - Partnerships
  - Assessment literacy?
- Empirical?
  - Documenting and evaluating program efficacy
- Policy
  - Accountability policy
  - Assessment policy

Resources

We have written extensively about balanced assessment systems including this paper well that you can download at: https://www.nciea.org/node/493