

Talking Points on *Balanced Assessment Systems: Trends in other states*
 Discussion with Legislative WASL Workgroup, Phone conference Oct. 13, 2008
 Brian Gong, Center for Assessment (Dover, NH)

1. Defining the purpose and use for any assessment is a key aspect—e.g., status and growth in accountability.

School Performance: Four Views		
(Carlson, 2001; Hill, 2001; Gong, 2002)	Status	Change
Achievement	“Status”: How high do students in this school score on state assessments?	“Improvement”: Is the performance of successive groups increasing from one year to the next?
Effectiveness	“Growth”: Are individual students learning as they progress from one grade to the next?	“Acceleration”: Is the school becoming <i>more</i> effective or improving more rapidly?

2. Assessment purpose and uses differ by level of governance and action: state, district, school, classroom, individual. (See previous discussion with Rick Stiggins.)
3. Calls for “balanced assessment systems” stem from:
 - Desire to combine multiple measures into overall decisions (e.g., accountability systems’ use of multiple indicators, weights, and combining rules)
 - Concerns that accountability systems have distorted desired focus for learning and teaching (e.g., counteract “narrowing of curriculum”)
 - Concerns that current assessment and accountability systems do not provide enough and the right information to support needed improvements in student learning and in school capacity (e.g., provide information to inform more local processes)

4. Aspects that people have advocated for in calling for “balanced, comprehensive, and coherent assessment systems”

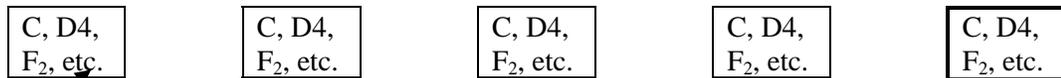
- Inclusive of students (especially different student characteristics)
- Inclusive of valued construct/content
- Provides “vertical” information for how to achieve goals or solve problems (e.g., summative, interim, formative assessment)
- Brings together in coherent way different actors who have different roles and responsibilities for achieving the goal
- System involves both assessing what the situation is and informing action (or acting) to improve the situation
- Informative of multiple purposes and uses

5. The purpose and use matters in assessment design

- Example: Design of interim assessments – Below are four test designs that differ in terms of the content included in a set of interim assessments administered four times during the year, followed by the state summative assessment.

Learning sequence of 10 topics/content standards during year									
A	B	C	D ₁₂₃₄	E	F ₁₂₃	G	H	I	J
Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June

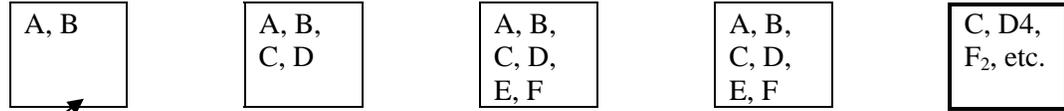
Four interim assessments & content topics assessed	State assessment & content assessed
--	-------------------------------------



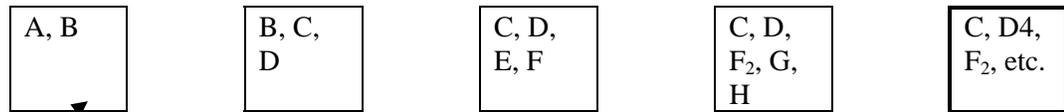
In this model, the interim assessment mirrors the end-of-year state assessment in terms of content, balance of emphasis, format, administration conditions, etc. Each test administered during the year covers the same content and has the same design. This design provides high “practice” and high “prediction” from the interim to the end-of-year state assessment.



In this model, the interim assessment focuses on the content that was instructed. Each interim assessment covers only the content in the most recent instructional period, and thus each test’s content differs from the others.



In this model, the interim assessment is designed to assesses what was instructed, but is cumulative, i.e., the assessment includes all topics instructed up to that point in time.



In this model, the interim assessment is designed to assess what was instructed, but is also cumulative for the topics that will be assessed on the state assessment.

6. What are opportunities to make state large-scale assessment systems more comprehensive, balanced, and beneficial for student learning?
 - Can make more comprehensive and balanced in some ways listed above (e.g., more content areas)
 - Cannot make traditional large-scale assessments “diagnostic”
 - Timing
 - Coverage
 - Standardization rules (e.g., on grade level)
 - Can get more program evaluation information
 - Technology (e.g., CAT) can help, but still quite limited

7. How can formative and diagnostic assessments be added to a “state assessment system”?
 - Different components
 - Aligned
 - Administered...
 - State’s role?
 - Provide models
 - Provide quality assurance
 - Provide policy support
 - Integrate with state large-scale system (standards, assessment scores, reports, accountability)
 - Fuel deep reform and school capacity
 - Other? – “theory of action”

8. Developing and Supporting Balanced and Coherent Assessment Systems: State policies and programs — Are other states beginning to re-think their state assessment systems? Which ones? What are they doing?

States have taken various approaches to supporting the development of balanced and coherent assessment systems.

Highly Centralized

- Develop and promulgate tight definition
- Establish state criteria and vetting process for interim assessments
- Require specific formative and interim assessment practices by schools identified through accountability system
- Tie state funds to local adoption of highly specified assessment uses
- State offers specific professional development around use of assessment information

Moderately Centralized

- Make state assessment data available in machine-readable formats (via state data warehouses or downloaded/CDs) so local education agencies can integrate with local assessment and other data
- State develops list of recognized vendors or establishes selective partnerships with professional development/technical assistance providers around assessment

Highly Localized

- Provide funds for local education agencies to engage in assessment professional development selected by LEA