Wyoming Assessment Task Force: Introductory Meeting

Joseph Martineau & Scott Marion, Center for Assessment

Brent Young & Deb Lindsey, WDE

Casper, WY

Center for Assessment

June 1, 2015

Overview of Day

- Introductions and group processes
- A shared understanding of key assessment concepts
- Goals, purposes, and uses of assessment data
- Thoughts about a comprehensive assessment system
- Outline of our final report

- We have designed an **agenda** that requires your **active** and **thoughtful participation**

Introductions

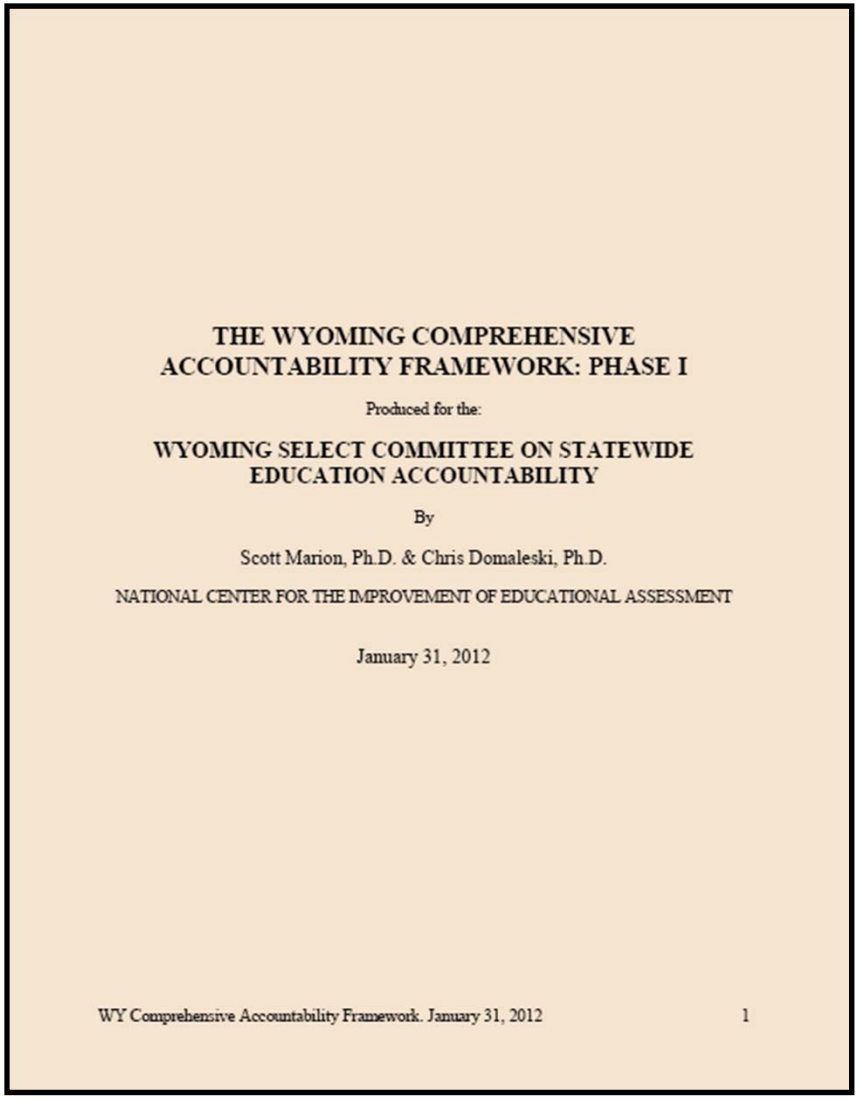
- Please tell us:
  - Your name
  - Where you live
  - Your role in education
  - Why you applied to be on the assessment task force

- We will then share a little background on the Center for Assessment

National Center for the Improvement of Educational Assessment (aka The Center for Assessment)

- Non-profit consulting firm established in 1998 with the mission of improving student learning through improved assessment and accountability practices
- Current contracts with 30+ states/entities
  - Almost all are long-term contracts designed to provide technical and design support for a range of assessment and accountability issues
- Purposely small—16 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.
- Independent of any governmental agency or testing company

Our Work in Wyoming Since 2011



THE WYOMING COMPREHENSIVE ACCOUNTABILITY FRAMEWORK: PHASE I

Produced for the:

WYOMING SELECT COMMITTEE ON STATEWIDE EDUCATION ACCOUNTABILITY

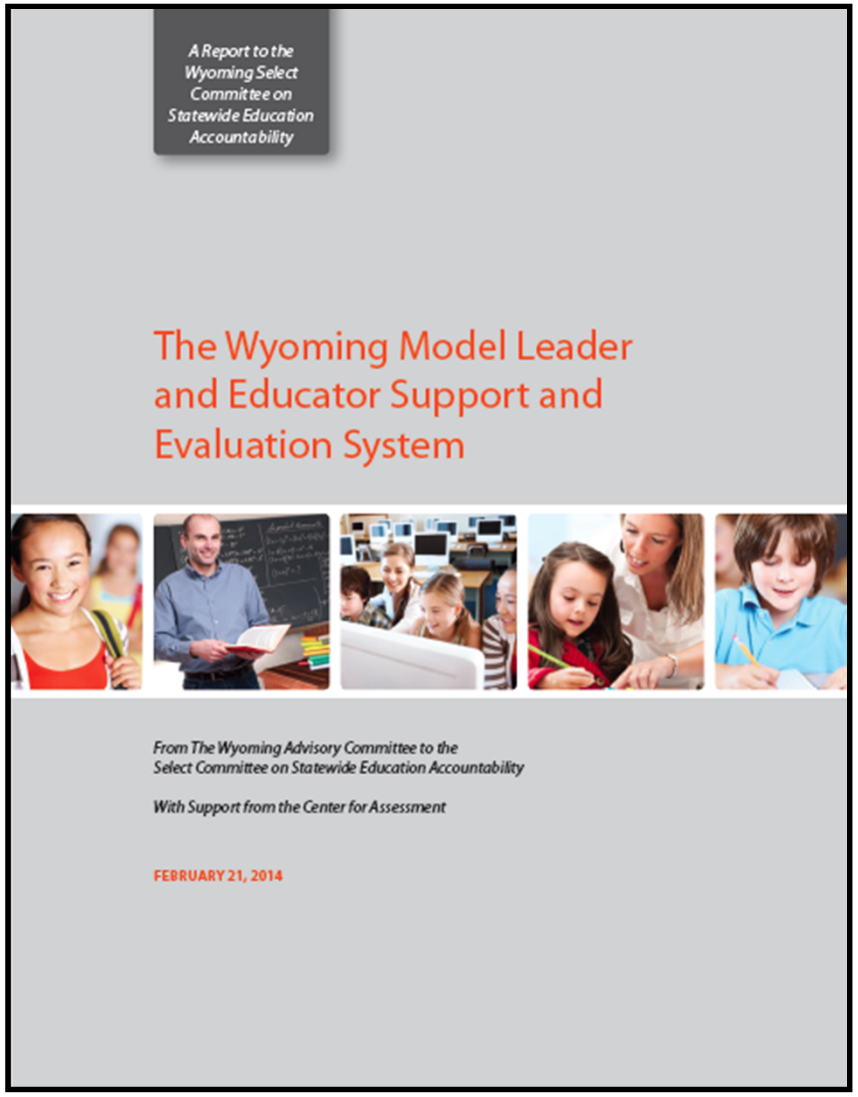
By

Scott Marion, Ph.D. & Chris Domaleski, Ph.D.

NATIONAL CENTER FOR THE IMPROVEMENT OF EDUCATIONAL ASSESSMENT

January 31, 2012

WY Comprehensive Accountability Framework. January 31, 2012 1



A Report to the Wyoming Select Committee on Statewide Education Accountability

The Wyoming Model Leader and Educator Support and Evaluation System

From The Wyoming Advisory Committee to the Select Committee on Statewide Education Accountability

With Support from the Center for Assessment

FEBRUARY 21, 2014

Center for Assessment

Our Role: Facilitate the Task Force

- The Center is under contract to the legislature, but works closely with WDE on a variety of issues…

(iii) *As authorized under section 7(b) of this act, the legislative service office, through acquired* ***professional consulting expertise****, and the* ***department of education****, shall assist the state board in its review and evaluation required by this subsection*;

- Joseph Martineau—Senior Associate at the Center since January 2015. 10-year career at MI DOE from psychometrician through deputy superintendent.
- Scott Marion—Associate Director (soon Exec Director) at the Center since 2003. Formerly, assessment and accountability director at WDE 1999-2003.

Charge to the State Board of Education: Statewide Assessment Review

Section 6.

(a) The state board shall conduct a review and evaluation of the statewide assessment system established by W.S. 21-2-304(a)(v) which not only fulfills the requirements of the **statewide education accountability system** prescribed under W.S. 21-2-204 and complies with statewide assessment and accreditation requirements imposed upon the state board under W.S. 21-2-304, but **strives for a high quality, rigorous and effective assessment adhering to principles of sound education policy and test measurement, with due consideration given to cost, testing time requirements for students and assessment burdens placed upon school districts**. The study shall **review assessment options available to the state and shall recommend an approach for the state in continuance of a statewide assessment system** required by law. In implementing this subsection, the following apply:

Role and Composition of Task Force

- (i) The state board shall assemble a task force to assist with the assessment review and evaluation. The task force shall be comprised of representatives of small and large school districts and schools from all geographic regions of the state and shall at minimum include representatives from district and school administration, school district assessment and curriculum program administrators, elementary and secondary school teachers, school district board members, state higher education representatives, member of the Wyoming business community and parents of children enrolled in Wyoming public schools;

Task Force Members

- 26 members select from a fairly large pool of applicants in the following categories:
  - School board members (2) and Superintendents (2)
  - Curriculum/assessment directors (4)
  - Principals (3)
  - Teachers: secondary (2), elementary (2), and Special populations (1)
  - Post-secondary education (2)
  - Parents (2)
  - Business (2)
  - Other (4)

Charge to the Task Force:
Recommend an Approach to Assessment & Transition

- (iv) On or before October 15, 2015, the state board shall report to the joint education interim committee and the select committee on findings and recommendations developed from its review and evaluation conducted under this subsection. **In addition to recommending an approach for the future statewide assessment system, recommendations shall provide necessary mechanisms and processes to support the transition from the statewide assessment system** existing on the effective date of this act to the student assessment developed and recommended under this subsection.

Structure of Task Force Work

- Center for Assessment and WDE have worked closely to develop a structure for supporting the Task Force in meeting their charge
- Three (3) in-person meetings
  - June 1
  - July 28-29
  - September 9
- Three (3) half-day webinar meetings
  - June 29
  - July 13
  - August 18
- The Center and WDE will…
  - Prepare materials and activities for each meeting
  - Draft sections of the final report based on Task Force work after each meeting (for review at the next meeting)

Purpose of This Meeting

- This Task Force kickoff meeting is intended to:
  - Ground members in an understanding of the different types of assessment as well as various technical issues associated with assessment design, use, and validation
  - Engage members in wrestling with what they really want from an assessment and/or assessment system
    - Assessment design is a good example of striving for optimization under often significant constraints and requirements

Goals For the Final Report

- Reflect the deliberations and recommendations of the Task Force
- Provide a solid framework to support the development of new legislation and/or amendments to existing statutes regarding student assessment
- Provide a solid framework for WDE in drafting one or more Requests for Proposals (RFPs) to provide requirements for vendors' bids to supply Wyoming's next assessment system

- Questions? Comments?

Task Force Authority

- In Scope of Authority
  - **Recommend** purposes and uses of a state assessment (system)
  - **Recommend** design characteristics that a state assessment (system) should have
  - Make **recommendations** for the process and supports needed to assure a successful transition to Wyoming's next assessment (system)
  - **Determine** whether to address a system of assessments, a set of assessments, or just the assessment given to the general population of students
- Out of Scope of Authority
  - **Select or identify** a new state assessment (system)
  - **Determine** the purposes and uses of the state assessment (system)
  - **Determine** design characteristics of the state assessment (system)
  - **Determine** the process and supports to be put in place for the transition

Establishing Group Norms

- **Group norms** (how we treat each other)
  - Treat each other with respect
  - Dissenting views should receive a hearing
  - Disagreement is with ideas, not people (for the one disagreeing)
  - Disagreement is not personal (for the one being challenged)
  - Trust the process and each other (but please provide feedback on the process)
  - Maintain confidentiality (i.e., don't broadcast partially-formed ideas as if they are a "done deal")
  - Other group norms that need to be added, any that need to be modified?

Establishing Group Rules

- A framework for how we get the work done:
  - Large-group facilitator role (Center for Assessment staff)
    - Manage time and completion of tasks on the agenda
    - Manage turn taking
    - Attempt to draw out less vocal members
  - Small-group facilitator role (selected by the small group)
    - Manage time and completion of tasks
    - Manage turn taking and inclusion of all group members
    - Participate in group tasks
  - Small-group note-taker/spokesperson role (selected by the small group)
    - Keep a record of the results of group tasks
    - Report out the results of the group tasks
    - Ask group members for clarification as needed in report out

Establishing Group Rules, continued…

- Large-group presentation expectations
  - In general, presenters expect that questions and comments will be brought up **throughout the presentation** rather than just at the end
  - Questions asking a presenter to clarify meaning may always be asked at any time
- Large-group turn-taking rules
  - Speakers will be recognized in order of request to a facilitator
  - If a task force member has a comment on or question **directly related** to the current speaker's question or comment, he or she may ask to speak before the focus of the discussion shifts
- Small-group turn-taking rules
  - Managed by the small group facilitator

Establishing Group Rules, continued…

- Managing possible conflict of interest
  - If a participant (including facilitators) has a possible conflict of interest on any topic on a meeting's agenda, that participant will notify a facilitators in advance of the meeting to discuss potential approaches to managing the conflict
- Encouraging free expression
  - All recommendation will be attributed to the task force as a whole (or to an unnamed majority or minority of the task force)
  - Task force members agree that task force activities may be described, but no information will be shared that may identify the views of a specific task force member
  - Task force members agree that the types of deliberations being undertaken by the task force may be discussed outside the task force, but potential recommendations will not be shared
  - Drafts of recommendations to be kept confidential until the report is completed and delivered

Establishing Group Rules, continued…

- Developing recommendations
  - Facilitators will attempt to briefly summarize recommendations at the end of a discussion
  - After the meeting, facilitators and WDE staff will draft fleshed out recommendations from discussion records
  - Order of decision-making
    1. Task force members will attempt to achieve consensus on summary recommendations
    2. Failing consensus, a supermajority (75%) of task force members will result in a single recommendation
    3. Failing a supermajority, facilitators will summarize a majority view (if applicable) and one or more minority views
  - Reviewing recommendations
    - Facilitators will provide drafts in advance of the next meeting for task force review
    - Task force members will provide feedback on the accuracy of drafts delivered by facilitators in a later meeting as provided for in the meeting agenda

Establishing Group Norms and Rules, continued…

- Norms and rules may be revisited by the task force at any time as needed

Break

- We start again at 10:00 AM
- September 9th

What We Know and Want to Know: Part 1

Part 1: 35-40 minutes *(end at least by **10:45**)*

- Count off by fours
- Label one sheet of butcher paper "**what we think we know**" and a second one, "**what we want to know**."
- Spend **15 minutes** listing the concepts "**what we think we know**." Discuss with your group **how** you know this before listing it on this sheet.
- Next, list the concepts that you think **you need to know more about** before providing advice on Wyoming assessment (15-20 minutes).
- We are providing you with a "**starter list**" of some measurement and assessment concepts (see handout) to consider as you do your work, but feel free to add more and/or ignore any of those listed.

What We Know and Want to Know: Part 2

*Part 2: 20-25 minutes (end by **11:05 AM**)*

- Each group will **post** its **responses** in the designated location on the wall.
- We will have an opportunity to **ask questions** about what other groups posted to make sure we have a shared understanding of key assessment topics.
- The facilitators will look to **expand on the various topics** and highlight topics for further discussion.

Goals for An Assessment (System)

- Assessment design is always a case of **optimization under constraints** (thanks to Henry Braun)
- The Wyoming assessment system **must** support both state and federal **accountability** requirements
- Any given assessment can serve **only** a **very limited number of purposes** well
- Assessments must be **validated** for **specific purposes**
  - For example, just because an assessment may be valid for one purpose (e.g., predicting college readiness, measuring student growth), it does not hold true that it is valid for any other purpose (e.g., informing instruction, measuring achievement against the state's content standards).

Goals for An Assessment (System): Part 1

- *Spend 5 minutes* ***working alone***, write down your highest priority purposes and uses for assessment results. In other words, what uses do you want the assessment to be able to support?
- *Spend 10 minutes in* ***groups of 3-4*** from those sitting around you and discuss what each of you has written down to search for commonalities. See if you can **eliminate** any purposes and uses.
- *Spend 5 minutes having a group note-taker* write your group's specific **highest priority purposes** on the butcher paper with the appropriate general category of use (e.g., program evaluation, instructional feedback
- *Spend 5 minutes* ***independently*** **reviewing** the work of the other groups.
- **End by 11:30 AM**

Goals for An Assessment (System): Part 2

- *5-minute **large group** discussion* to ask clarifying questions of other groups
- *Spend 10 minutes **independently*** identifying your top three priorities from all the priorities listed
  - Use the markers to write next to your top three, a **1=highest priority**, **2= 2nd highest priority**, and a **3=3rd highest priority**
- *15-minute group discussion and debrief*. Do we want to make any changes to the emerging group priorities?
  - Note: this is just a draft at this point, you will have a chance to review these decisions at the next meeting
- We are providing you a handout to keep track of your task
- **End by 12:15**

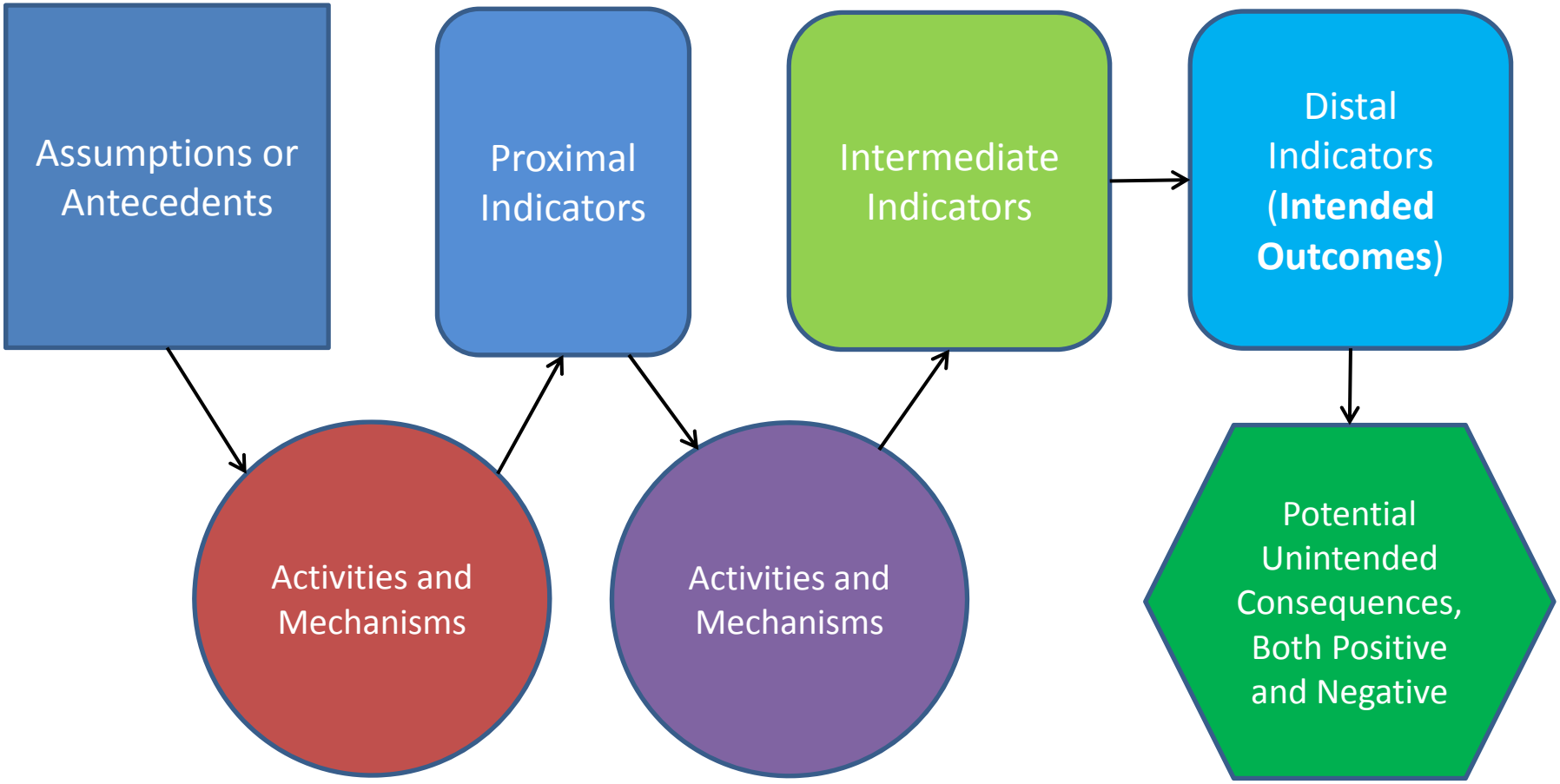
Lunch

- We start again at 1:00 PM

Purposes and Uses (15 minute intro)

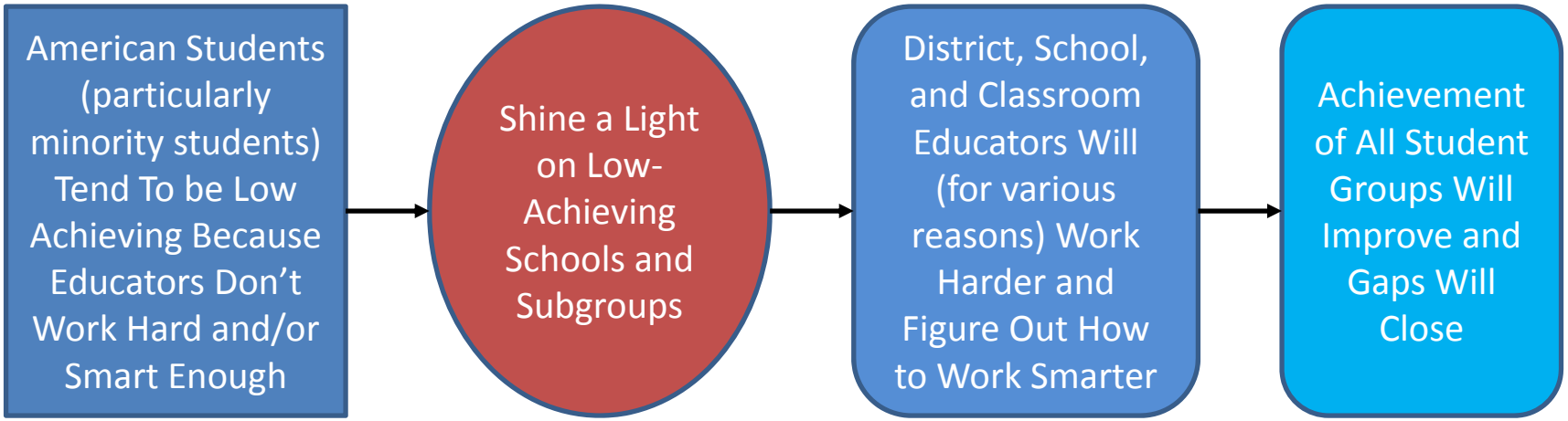
- We suspected pretty strong agreement on many of the big-picture purposes and uses
- Now we get down to the **details**…
- It's easy to say that we want an assessment to do X, Y, and Z, but much harder to say **how** we expect the assessment results to work within an educational system to ensure that X, Y, and Z get done
- We have found **theories of action** or logic models to be useful tools to help think through these issues

Basic Structure of a Theory of Action



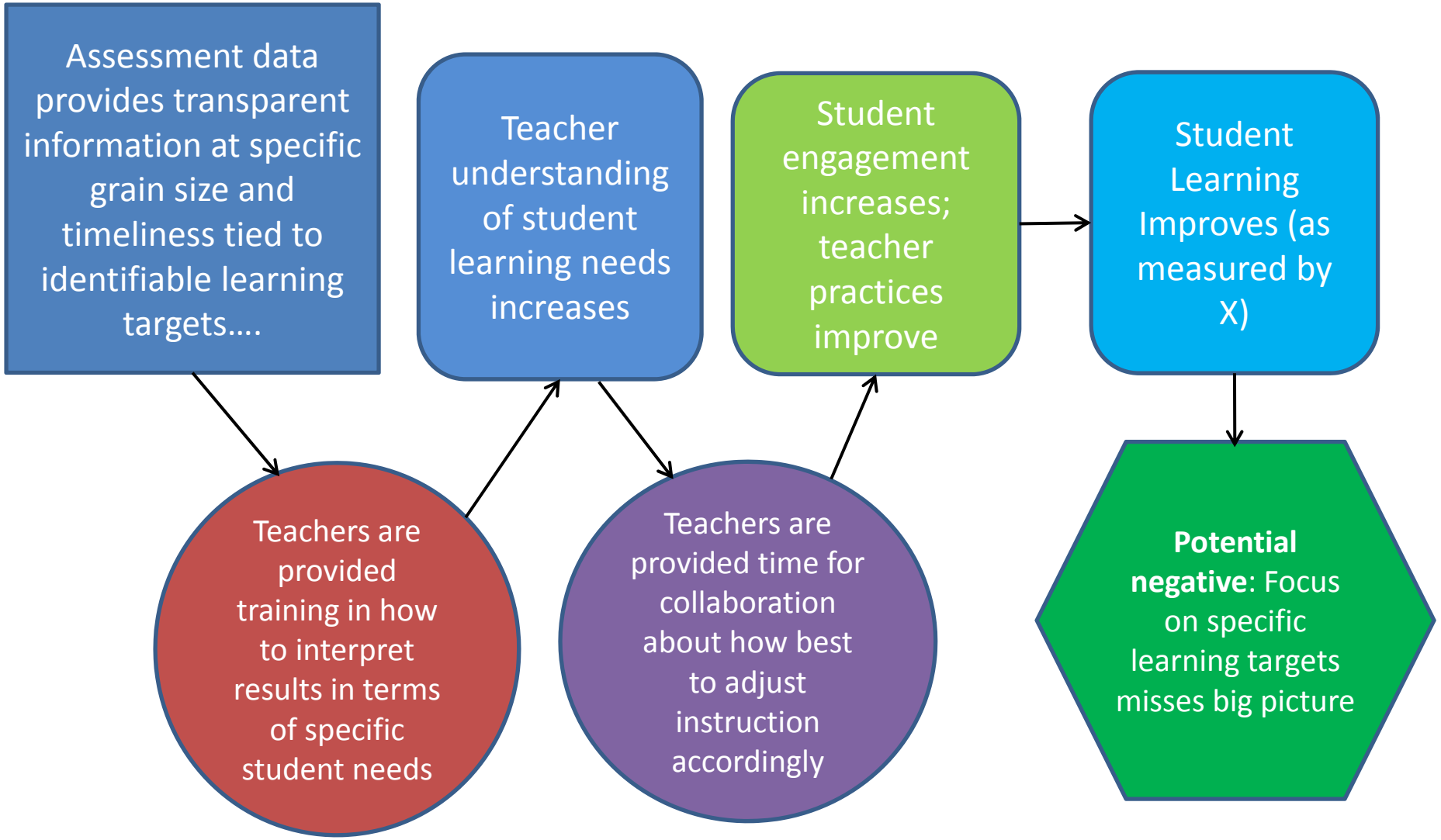
An Example of an Implicit Theory of Action

- No Child Left Behind (the theory of action was never articulated, but can be inferred from the law's structure)



- Missing a few pieces
- Since the passage of NCLB, there has been much work on developing a theory of action to support education reform so that we don't end up with an implicit theory of action that is unlikely to be successful.

An Example: Assessment for Learning



Purposes and Uses

- While we do not expect you to draw extensive theory of action diagrams (but you can!), we do want you to **engage** in the process of **thinking** through **how** you will go **from the assessment to realizing your goals**
  - In other words, **how will the assessment results support** your intended uses and purposes
  - We will **press you to be as specific** as possible, probably annoyingly so

Purposes and Uses

- **Count off** by sixes to form groups. We will need a volunteer to take careful **electronic notes** for each group.
- As a group, **quickly** select **one** of the **big picture goals** as a starting place. We will ***encourage*** groups to select certain goals to try to make sure all are covered by at least one or two groups.
- Try to add **specific details** to better define this goal and ensure that all group members have a shared understanding of this goal.
- At a high level, **how** do you envision **using** assessment data to help you realize this goal?
  - For example, your group might say, "we intend to use these data to help us evaluate the degree to which various curricula and/programs are working as intended."

Purposes and Uses

- Clearly articulate **how** these **data** will be used to **achieve your goals**. The following questions might be useful:
  a. Describe as specifically as possible the nature of the data you need to support your intended uses.
  b. How would the results need to be reported to support your intended uses?
  c. How often during or across years would you need these data?
  d. What "**claims**" would you like to be able to make on the basis of these assessment results?
  e. How **quickly** after the assessment(s) is/are completed would you need the data? If you say "immediately," please specify how the results will be used <u>immediately</u>!

Purposes and Uses

d. How would these data need to be transformed, if at all, to be useful (e.g., transforming 2 or more test scores into growth calculations)?

e. Would the assessment data need to be tied to specific curricula or learning progressions to serve your uses or can it be tied to broader content standards?

f. What knowledge and skills would teachers need to use these data to support your intended uses?

g. If they do not possess these skills now, how will they acquire them?

h. What are the potential unintended negative consequences of using the assessment results as you describe? If you cannot think of any, think harder!

Purposes and Uses Summary

- Once you work through your logic model, **please summarize your results** (diagrams are acceptable!) and nominate a presenter.
- We will spend the next 45-60 minutes reviewing the results of the groups' work, asking questions, and probing each group's logic models.

- Email to jmartineau@nciea.org
- smarion@nciea.org

Comprehensive Assessment Systems

- You get a brief rest!
- We will provide a ±30 minute presentation about comprehensive assessment systems
- This will be followed by a large-group discussion to help provide us with direction as we conceptualize plans for the rest of the Task Force work

Center for Assessment

Can you handle the truth?

- **End-of-year**, **statewide summative assessments** will **NEVER be useful to inform (adjust) instruction**
- **Why?**

- If well-constructed, such assessments can support many types of **evaluative** activities
- To be clear, curriculum and programs can be adjusted based on those evaluations, but this is an after-the-fact use
- Large-scale assessments can also provide meaningful data for school and perhaps educator accountability systems
- Remember, the assessment(s) **MUST support school accountability uses**

What are our options?

- **Comprehensive assessment systems** (or at least notions of these systems) have emerged as a potential solution
- If each individual assessment can serve only a limited number of purposes, the thought (hope) is that a **comprehensive** assessment **system** can serve **multiple** and varied **purposes**
- Except that there is **little agreement** on what turns a set of assessments into **a comprehensive system**
- *A collection of assessments does not entail a system any more than a pile of bricks entails a house* (Coladarci, 2001).

Comprehensive Assessment Systems

- **Multiple types** of assessments—**summative**, **interim**, **formative**
  - Do all have to be present for an assessment system to be comprehensive?
- **Multiple loci of control**—classroom, school, district, and state
  - Interim and summative can be found at all levels of the system
  - Formative is only found at the classroom level
- What is the role of **local curriculum** (not standards) in assessment design?
- Can a "**state**" implement a comprehensive assessment system in a **local control** context?
  - If so, how? If not, what are our options?

Comprehensive Assessment Systems

- For assessment systems to be "**comprehensive**," the various pieces must be **coherently designed to work together**
- That sounds good, right?
- How do we **ensure coherence** among multiple and diverse components?
- Fortunately, we are not starting from scratch…

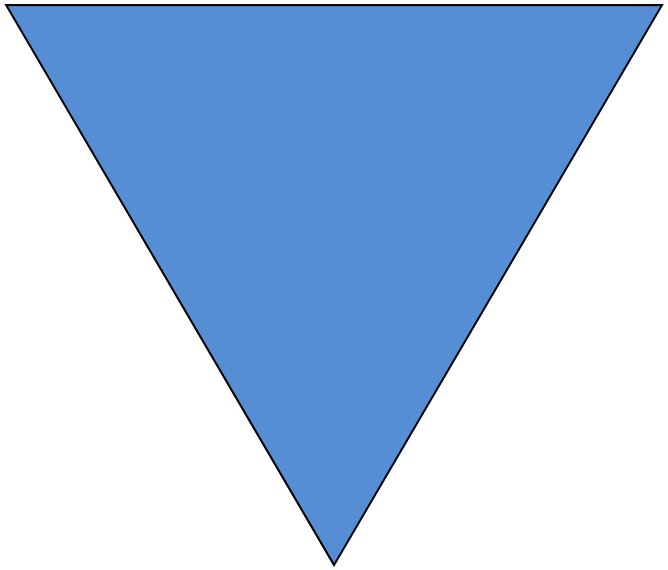
Knowing What Students Know

- *Knowing What Students Know*: *The science and design of educational assessment* (NRC, 2001) synthesized a tremendous body of learning and measurement research
- It set an ambitious direction for the development of more valid assessments and assessment systems
- It helped make many of the concepts in Mislevy's Evidence-Centered Design more understandable

The Assessment Triangle

Observation

Interpretation



Cognition

A Heuristic

- The assessment triangle is a heuristic to help **guide recommendations about design** and **organize an evaluation of validity**
- The triangle gives an **important framework** to guide our work should we pursue a comprehensive assessment system
  - Any assessment or assessment system must ensure coherence among:
    - The learning model/expectations (cognition)
    - How we collect the assessment data (observation)
    - How we interpret and report those data (interpretation)
  - Therefore, the design of a coherent and comprehensive assessment system must be **based on a common conception of learning** (more than standards)
  - Whether or not we pursue an assessment system, we will revisit this idea of coherence throughout the task force work

Comprehensive Assessment Systems

- Therefore, designing a comprehensive assessment system is more like creating a menu that works as a whole for a fine dining experience rather than making one's way through an extensive buffet line
- In doing so, we need to think about **how the information flows and joins together** across components to maximize the following three things:
  - Available information to inform the way each component is implemented
  - Information the system as a whole makes available to students, educators, and policymakers
  - The coherence of the information (so that as a whole it tells an **accurate and clear story** about students, classrooms, schools, districts, and the state)
- **This is hard**! But it is especially relevant if we aspire to use any local information for accountability purposes

Comprehensive Assessment Systems

Your turn again!

Spend about 5-10 minutes writing brief **individual** responses to the following questions:

1. What benefits does a comprehensive assessment system offer for WY specifically?
2. What are some logistical, technical, and policy challenges associated sound implementation in WY?
3. Should we explore a comprehensive assessment system design as part of the Task Force's work?

Let's spend about 15 minutes in a large group discussion of your responses….

Final Report Discussion

- ***Very Rough Draft*** Table of Contents
  I. Executive Summary
  II. Recommended Purposes and Uses of Assessment, and Intended Outcomes of Implementation
  III. A Wyoming Assessment System
  IV. Design Considerations for Types of Assessments and Tasks in a Coherent Assessment System
  V. Recommendations for Implementation
  VI. References/Sources Consulted
  VII. Appendices as Necessary

Very Rough Ideas for Chapter Contents

II. Recommended Purposes and Uses of Assessment, and Intended Outcomes of Implementation

a. Specific purposes and uses of assessment system and specific assessments
b. Intended outcomes
c. Theory of action for achieving intended outcomes
d. Necessary conditions for implementing theory of action (systems approach)
e. Appropriate and inappropriate high-stakes uses

Very Rough Ideas for Chapter Contents

III. A Wyoming Assessment System

a. Rationale for a comprehensive assessment system
b. State-level components of a comprehensive assessment system and uses served
c. District/school-level components of a comprehensive assessment system and uses served
d. Coordination and coherence among the state and local components

Very Rough Ideas for Chapter Contents

IV. Design Considerations for Types of Assessments and Tasks In a Coherent Assessment System

a. Introduction to *Evidence Centered Design* and *Knowing What Students Know*
b. The role/capacity of various item and task types to provide necessary evidence for specific uses
c. Design considerations and characteristics of various types of assessment
d. Summative (computer-based testing, common items, sampling, adaptive?)
e. Interim (predictive, instructional, evaluative, looking back, looking forward, pacing?)
f. Formative (tools, professional development, mentoring?)
g. Connections among various assessment types (summative, interim, and formative; predictive and evaluative; norm-referenced and criterion-referenced)
h. Links to necessary conditions for successful implementation of the theory of action

Very Rough Ideas for Chapter Contents

V. Recommendations for Implementation
   a. Timing
   b. Key requirements for an assessment RFP
   c. Mechanisms and processes to support effective implementation
   d. Link to necessary conditions for successful initial implementation of the theory of action, and maintenance over time
   e. Evaluating the validity and other technical qualities of the assessment system
   f. Evaluating the consequences of implementation

Next Steps

- Center for Assessment will **draft** the section of the report corresponding to this meeting, reflecting the deliberations of Task Force Members
- Task Force Members will receive the draft at least one week before the next meeting
- Task Force Members should carefully review the draft before the next meeting
- During the next meeting, we will do a big-picture review and feedback session regarding the draft, and will edit the draft based on the feedback
- Section II will be discussed at the next meeting
- We will also send you pre-reading materials for the next section before the next meeting

Thank You!

- We know that this is a heavy commitment of your valuable time!
- Therefore, we commit you that we will do everything we can to make sure this is a successful experience and that your voices are heard.
- Thank you on behalf of the State Board, the legislature, WDE, the Center, and the citizens of Wyoming.
- Happy Reading!