• Looking back on change in assessment and accountability
  • Center has made contributions; more to try to learn from past to inform future efforts
• Looking forward to future focus in assessment and uses
• A focus on validation and evaluation to improve

Refer to presentation slides and papers posted on www.nciea.org under “Annual Conference” here
Main points

• Major changes in assessment and accountability are/will happen; we must adapt to survive and contribute

• Better, more extensive validation and evaluation will be important key capacities in the future for all aspects of assessment and accountability (and other uses of assessment information)

• The Center is working to stay ahead of the curve; what else should we (and others) be doing?
Back to the Future: fierce change coming

• Phase I Pre-1994: **Norm-referenced testing**; focus on district market; dominated by three companies [CTB, Harcourt, Riverside]

• Phase 2 1994-2014: Shift to standards-based view of educational goals; wide-spread acceptance of performance [essay] assessment; passage of federal law; focus on individual states as responsible agent and market; rise of “**state custom assessment companies**” [Advanced Systems/Measured Progress, Pearson, DRC, Measurement Inc., Questar, ETS; CTB, PsychCorp/Harcourt]
• Phase 3 2014-present: affirmation of standards-based and state-centric models; some interest in more comprehensive assessment systems; ARRA consortia; focus on computer-administered tests; states sharing resources/contractors; necessity of company having a **scalable online platform**; trade-off CBT for performance assessments [Smarter Balanced, AIR, Pearson, DRC, ETS/Questar, Measured Progress, ACT, College Board, Measurement Inc., KU/DLM, NWEA, PARCC/New Meridian]
## Looking forward: Future shifts

<table>
<thead>
<tr>
<th>Current focus</th>
<th>Future focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment as is with current technologies</td>
<td>Assessment &quot;as is&quot; with emerging technologies</td>
</tr>
<tr>
<td>Attending to pieces</td>
<td>Attending to systems</td>
</tr>
<tr>
<td>Assessment and accountability issues primarily characterized as standardized measurement</td>
<td>Consider use in design and evaluation of measurement Measurement becomes support for primary use</td>
</tr>
<tr>
<td>Advocacy of “good ideas” and design</td>
<td>Technical detail that support quality implementation and Evaluation</td>
</tr>
<tr>
<td>Specialized, one-time studies by external agents to “establish” alignment, validity, effectiveness, impact, often for compliance</td>
<td>Embedded, continuous validation and self-evaluation, usually to inform improvement</td>
</tr>
<tr>
<td>State as primary unit of validation and evaluation activities</td>
<td>Many organizations (e.g., districts, user groups, consortia, states) sponsor and engage in validation and evaluation activities</td>
</tr>
</tbody>
</table>
Validation and Evaluation

• Validation ↔ Evaluation

• Each/both necessary to establish credibility, monitor usefulness, and inform efforts to improve

• Changes in endeavors may require validation/evaluation to stay constant in principle but different in application; or may require concomitant changes in validation/evaluation (assumptions, criteria, processes, agents, contextual understanding, etc.)
## Prior Emphasis vs. Shifted Emphasis

<table>
<thead>
<tr>
<th>Prior Emphasis</th>
<th>Shifted Emphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation of summative assessments in highly standardized contexts</td>
<td><strong>Supporting more flexible summative assessments</strong>; designing interim assessments and articulating the relationships between formative, interim, and summative assessments in comprehensive assessment systems</td>
</tr>
<tr>
<td>Attention to accountability systems’ “audience and purposes”</td>
<td>Developing tools to help define and evaluate theories of action; tools to use theories of action in program evaluations</td>
</tr>
<tr>
<td>Attention on design of accountability systems for identification of low-performing schools</td>
<td><strong>Design accountability systems that include systems of support</strong> and articulate a theory of action beyond “accountability ratings will motive schools to improve” or “schools will figure out how to improve”</td>
</tr>
<tr>
<td>Isolated technical advice or “one-off” reactive assistance</td>
<td>Developing tools that capture the underlying intelligence of the technical advice and make it possible for end-users to explore accountability systems by applying what-if reasoning with the tools</td>
</tr>
<tr>
<td>Expertise applied at the design stage of system development</td>
<td>Applying expertise and technical assistance to <strong>evaluation</strong> definitions, criteria, procedures, tools, and recognized, trusted authorities to support capacity building</td>
</tr>
<tr>
<td>Technical support on limited domains and measures</td>
<td>Supporting the examination, implementation, and validation of expanded domains and less traditional measures across systems</td>
</tr>
<tr>
<td>The development of separate assessment, accountability, and support systems</td>
<td><strong>Developing cohesive assessment and accountability programs</strong> that include support, curriculum, instruction, and other programmatic systems (e.g., career and technical education programs)</td>
</tr>
</tbody>
</table>

System coherence between “college” and “career” assessment, accountability, and support

• Erika Landl
The Future: What next? Accountability evaluation

Accountability Full System

Focus on design of rules

Focus on evaluation of performance to improve

Accountability classification only

Accountability evaluation: from identification to support, from design to evaluation

• Juan D’Brot
The Future: What next? Assessment validation

Assessment validation for more flexible assessment designs

• Leslie Keng
• 9:55 Break and transition to small groups
• 10:10 Small group discussions
  • Erika: Validation through college/career system coherence– Corner A
  • Juan: Validation of ESSA accountability systems – Corner B
  • Leslie: Evaluating assessment comparability – Corner C
• 11:00 Transition to large group
• 11:05 Validation/Evaluation Sessions: Discussants’ closing comments
• 11:45 Wrap-up/Concluding Thoughts
• 12:00 Lunch
For more information:

Center for Assessment

www.nciea.org

Brian Gong – bgong@nciea.org
Erika Landl – elandl@nciea.org
Juan D’Brot – jdbrot@nciea.org
Leslie Keng – lkeng@nciea.org