

*RILS 2018 Looking Back, Looking Forward:*  
Contributions and Challenges in Validation and Evaluation  
of Assessment and Accountability Systems

Brian Gong introduction to session by  
Erika Landl, Juan D'Brot, and Leslie Keng  
Center for Assessment

# Overview

- 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](http://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~~]

# Back to the Future: From paper to computer-based

- 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

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

# Validation and Evaluation

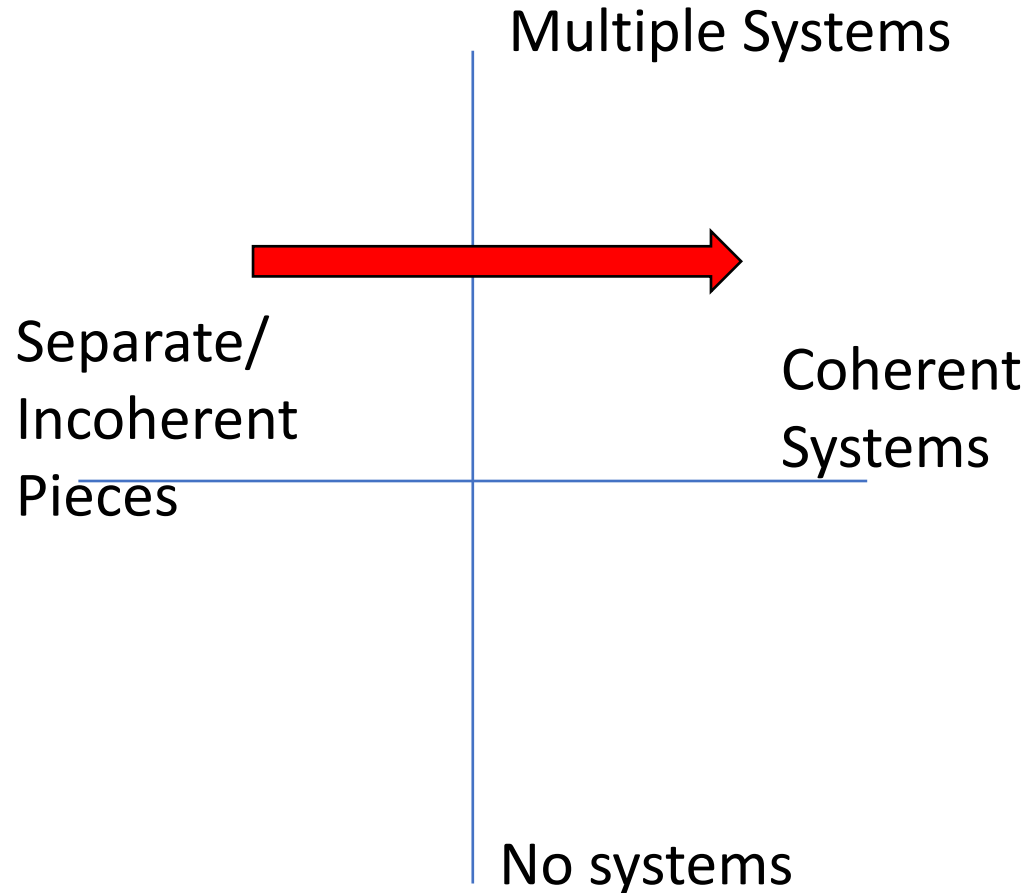
- Validation  $\leftrightarrow$  Evaluation
- Each/both necessary to establish credibility, monitor usefulness, and inform efforts to improve
- Changes in endeavors may require validation/evaluation to stay constant in principle but different in application; or may require concomitant changes in validation/evaluation (assumptions, criteria, processes, agents, contextual understanding, etc.)

# Looking forward: Future shifts for validation/evaluation

Prior Emphasis	Shifted Emphasis
Validation of summative assessments in highly standardized contexts	<b>Supporting more flexible summative assessments;</b> designing interim assessments and articulating the relationships between formative, interim, and summative assessments in comprehensive assessment systems
Attention to accountability systems' "audience and purposes"	Developing tools to help define and evaluate theories of action; tools to use theories of action in program evaluations
Attention on design of accountability systems for identification of low-performing schools	<b>Design accountability systems that include systems of support</b> and articulate a theory of action beyond "accountability ratings will motivate schools to improve" or "schools will figure out how to improve"
Isolated technical advice or "one-off" reactive assistance	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
Expertise applied at the design stage of system development	Applying expertise and technical assistance to <b>evaluation</b> definitions, criteria, procedures, tools, and recognized, trusted authorities to support capacity building
Technical support on limited domains and measures	Supporting the examination, implementation, and validation of expanded domains and less traditional measures across systems
The development of separate assessment, accountability, and support systems	<b>Developing cohesive assessment and accountability programs</b> that include support, curriculum, instruction, and other programmatic systems (e.g., career and technical education programs)



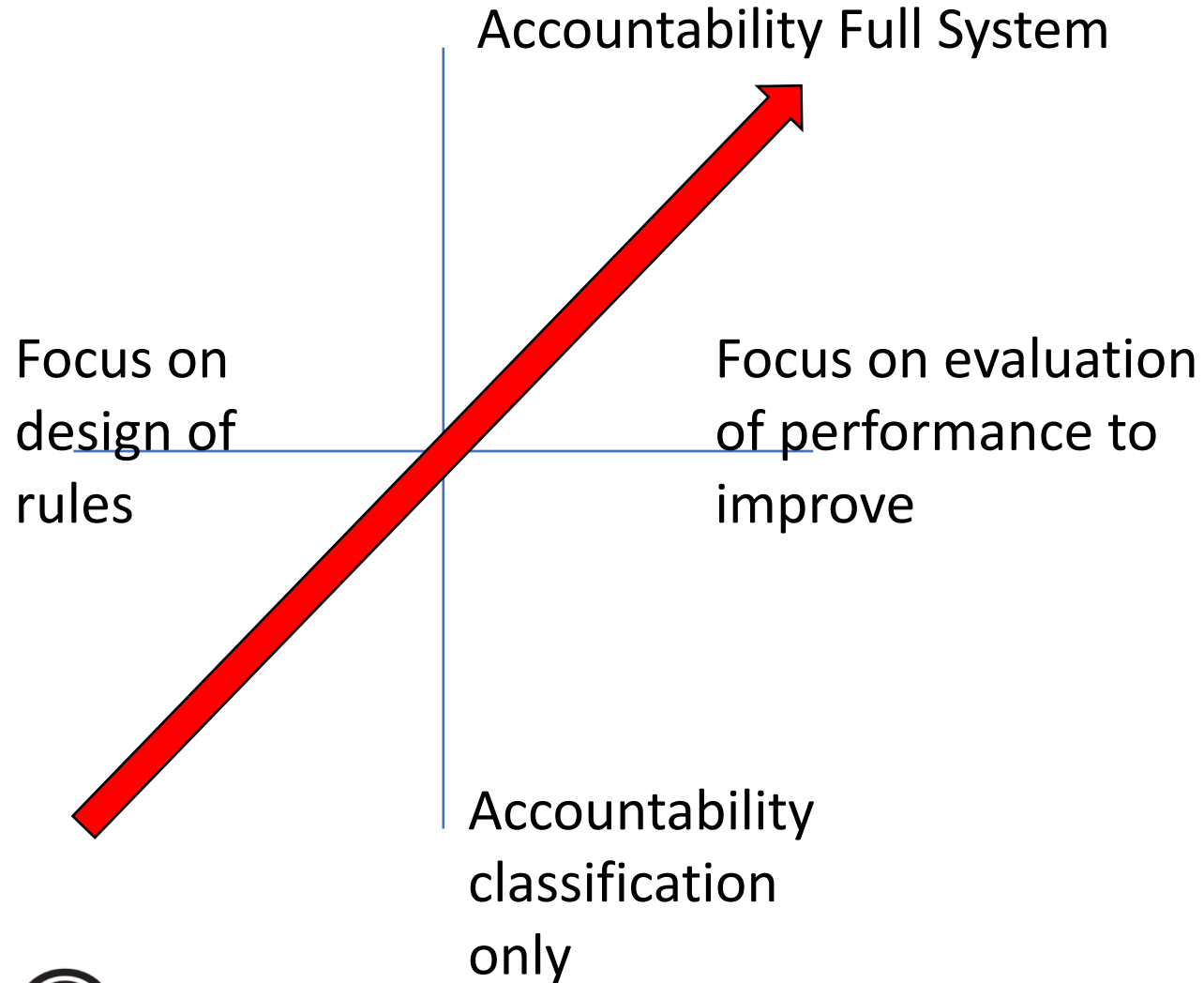
# The Future: What next? More coherent systems



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

- Erika Landl

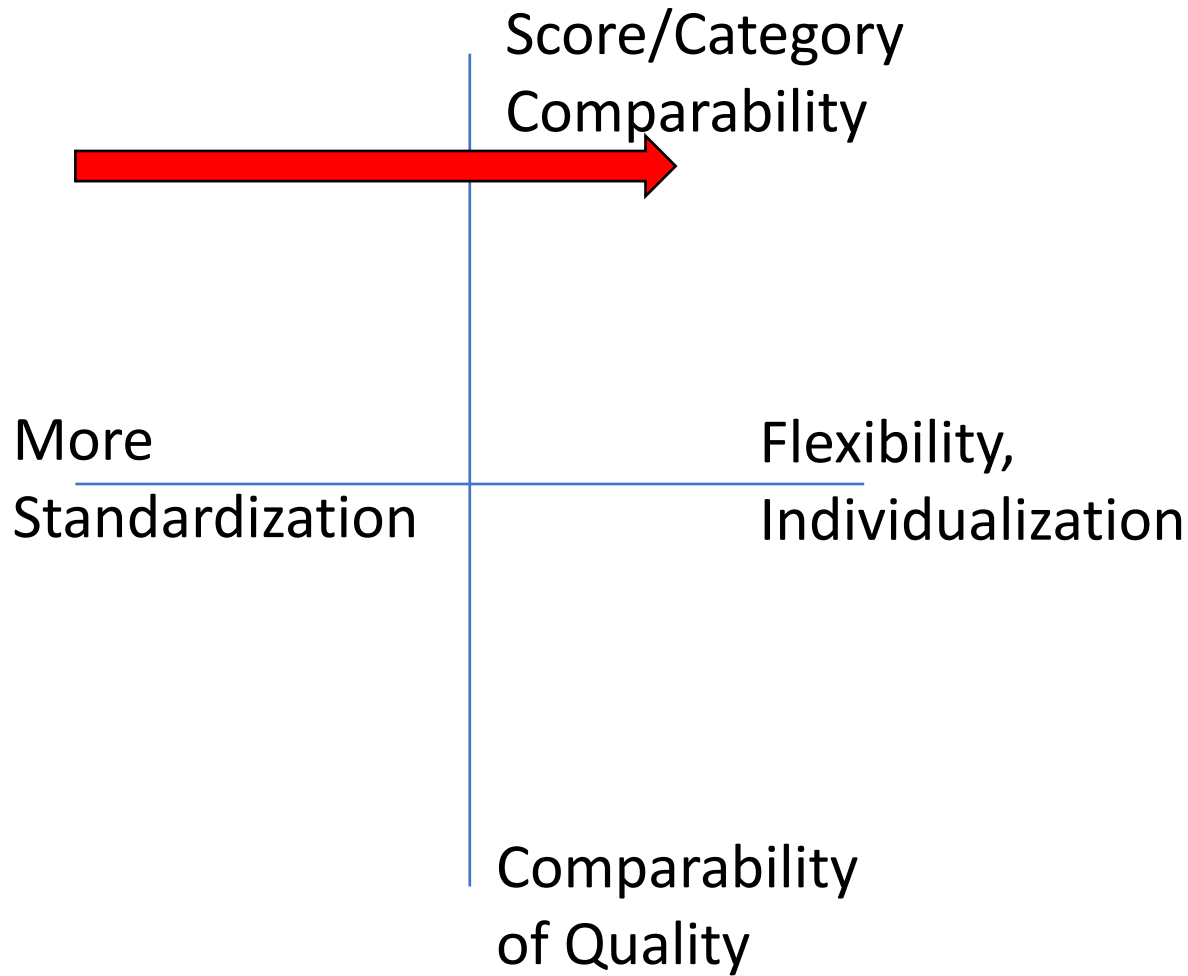
# The Future: What next? Accountability evaluation



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

# Logistics

- 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](http://www.nciea.org)



Brian Gong – [bgong@nciea.org](mailto:bgong@nciea.org)  
Erika Landl – [elandl@nciea.org](mailto:elandl@nciea.org)  
Juan D’Brot – [jdbrot@nciea.org](mailto:jdbrot@nciea.org)  
Leslie Keng – [lkeng@nciea.org](mailto:lkeng@nciea.org)