



# **RILS 2021: Design Innovation in Educational Assessment Systems**

*The National Center for the Improvement of Educational Assessment*

September 2021

# Welcome to RILS 2021

---

## Purpose of RILS 2021:

Discuss the different aspects of the design innovation process for assessment systems including:

- exploring the process for designing an innovative assessment system,
- providing an overview of current innovations in assessment systems, and
- delving into the attributes necessary for the design of an innovative assessment system.



# RILS Sessions

Session	Date	Presenter(s)
<b>Session 2:</b> Understanding the Problems for Design Innovation – <i>Root Cause Analysis</i>	Monday, September 20, 2021 1:00-2:30	Juan D’Brot and Chris Brandt Guests
<b>Session 3:</b> Leveraging Community for Design Innovation - <i>Engaging Stakeholders</i>	Monday, September 20, 2021 3:00-4:30	Carla Evans Guests
<b>Session 4:</b> Planning for Design Innovation – <i>Assessment Systems and Theory of Action</i>	Thursday, September 23, 2021 1:00-2:30	Erika Landl Nathan Dadey
<b>Session 5:</b> Exploring the Design Innovation Process – <i>Iteration in Assessment System Design</i>	Thursday, September 23, 2021 3:00-4:30	Brian Gong Guests
<b>Session 6:</b> Exploring the IADA Innovation Process – <i>Challenges and Opportunities</i>	Friday, September 24, 2021 1:00-2:30	Scott Marion and Carla Evans  Chris Domaleski

# Design Innovation – What is it?

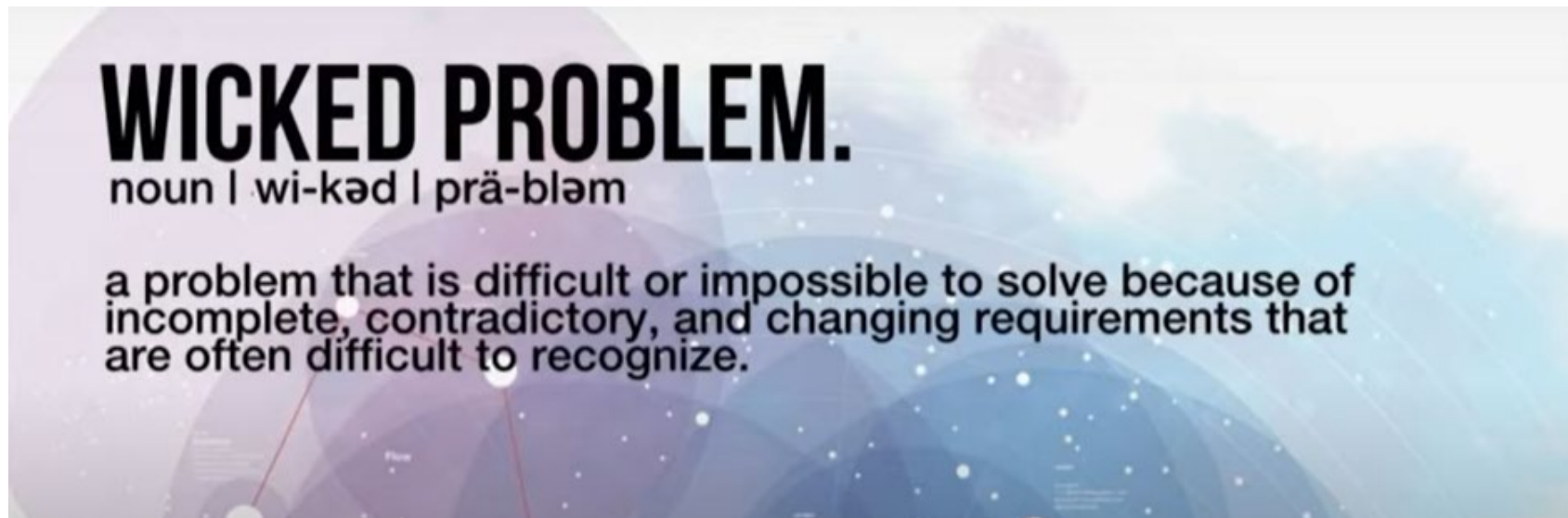
---

Ideology

+

Process

*To solve “wicked” problems in a user-centric way*



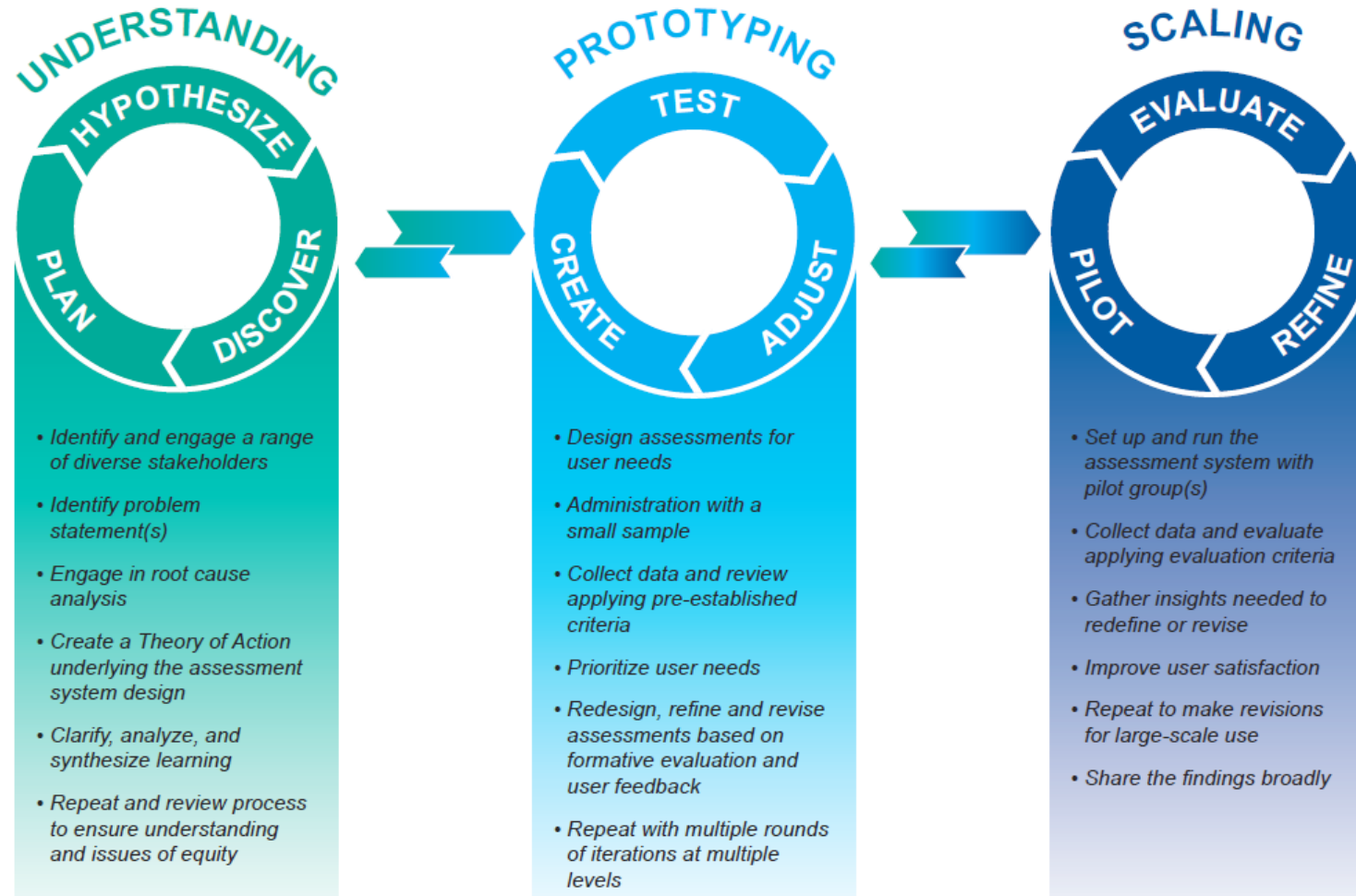
# Assessment Systems as a *Wicked Problem*

## *How to appropriately assess students?*

- *What does this mean?*
- *What is the end result*
- *What is the end result?*
- *Will it work for all students?*
- *Will it work in all content areas?*
- *Will it work in all environments?*
- *How does the solution of one aspect impact the next situation?*



# One Model for the DESIGN INNOVATION PROCESS



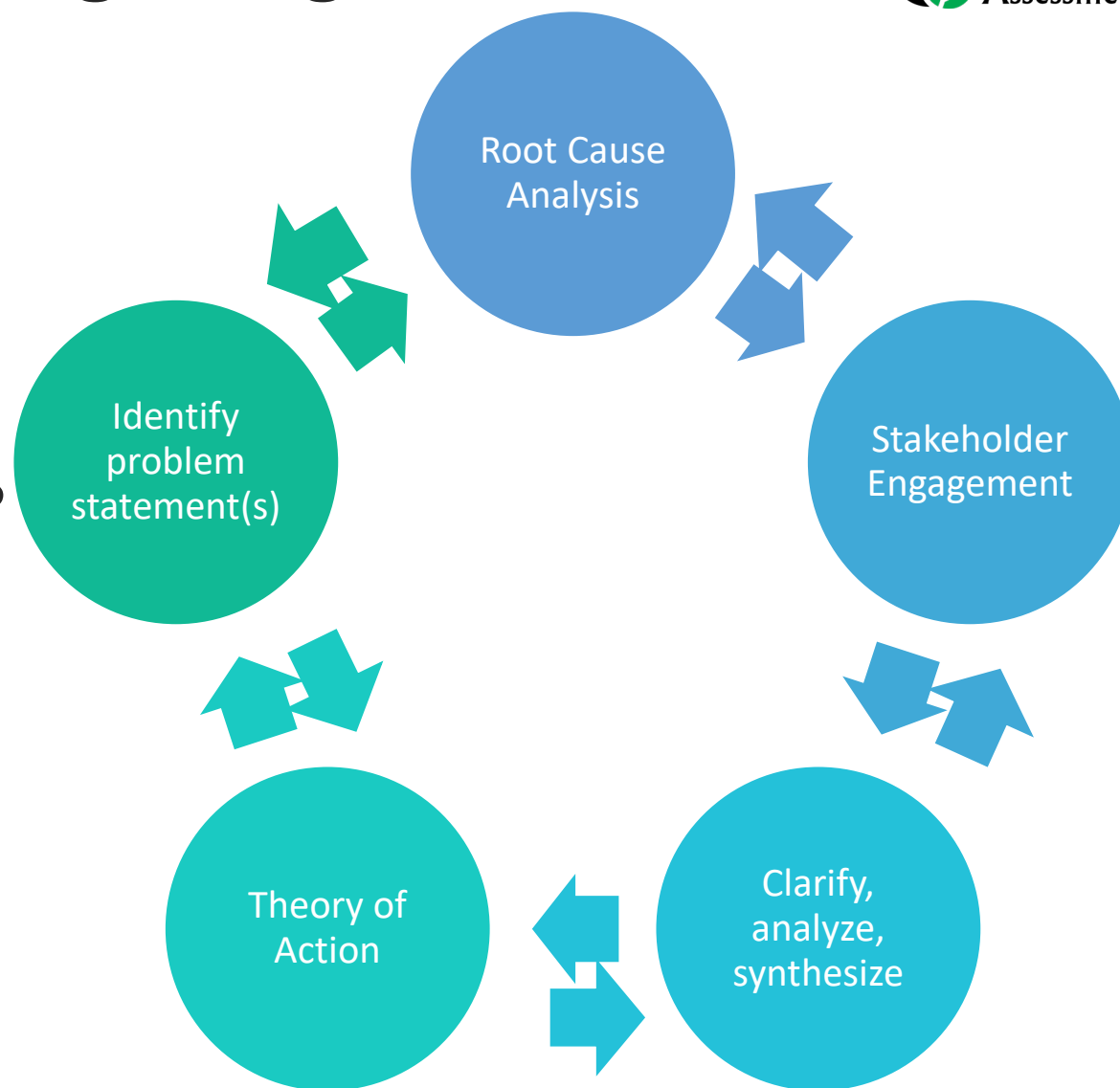


# “Understanding” Stage

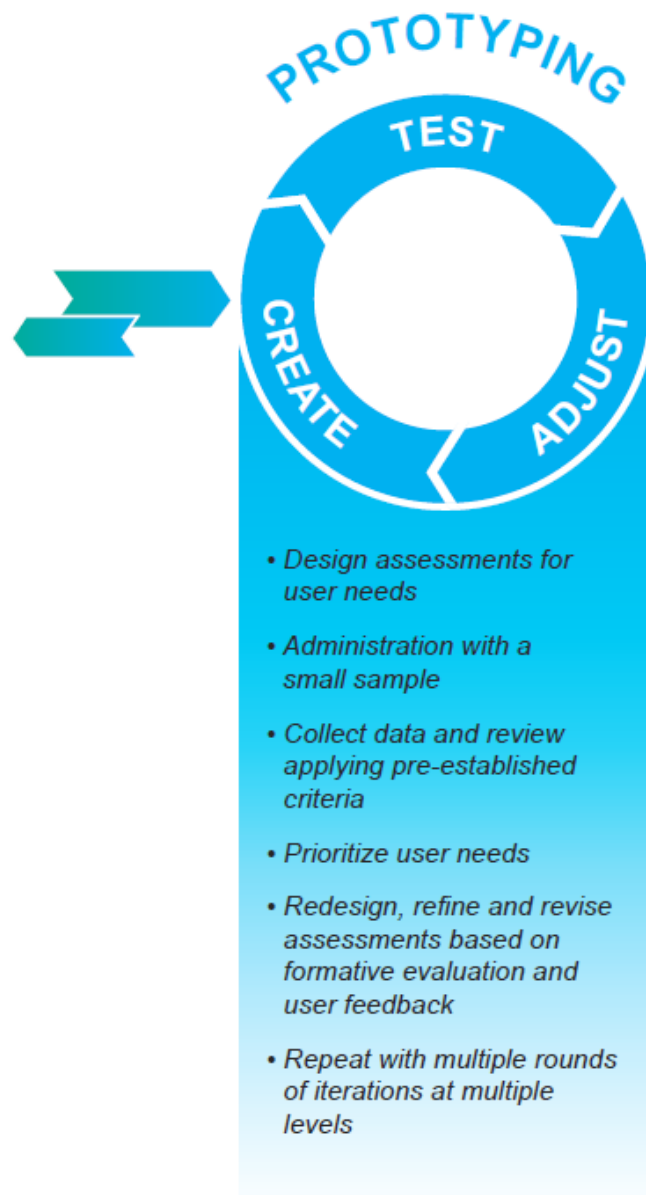


- Identify and engage a range of diverse stakeholders
- Identify problem statement(s)
- Engage in root cause analysis
- Create a Theory of Action underlying the assessment system design
- Clarify, analyze, and synthesize learning
- Repeat and review process to ensure understanding and issues of equity

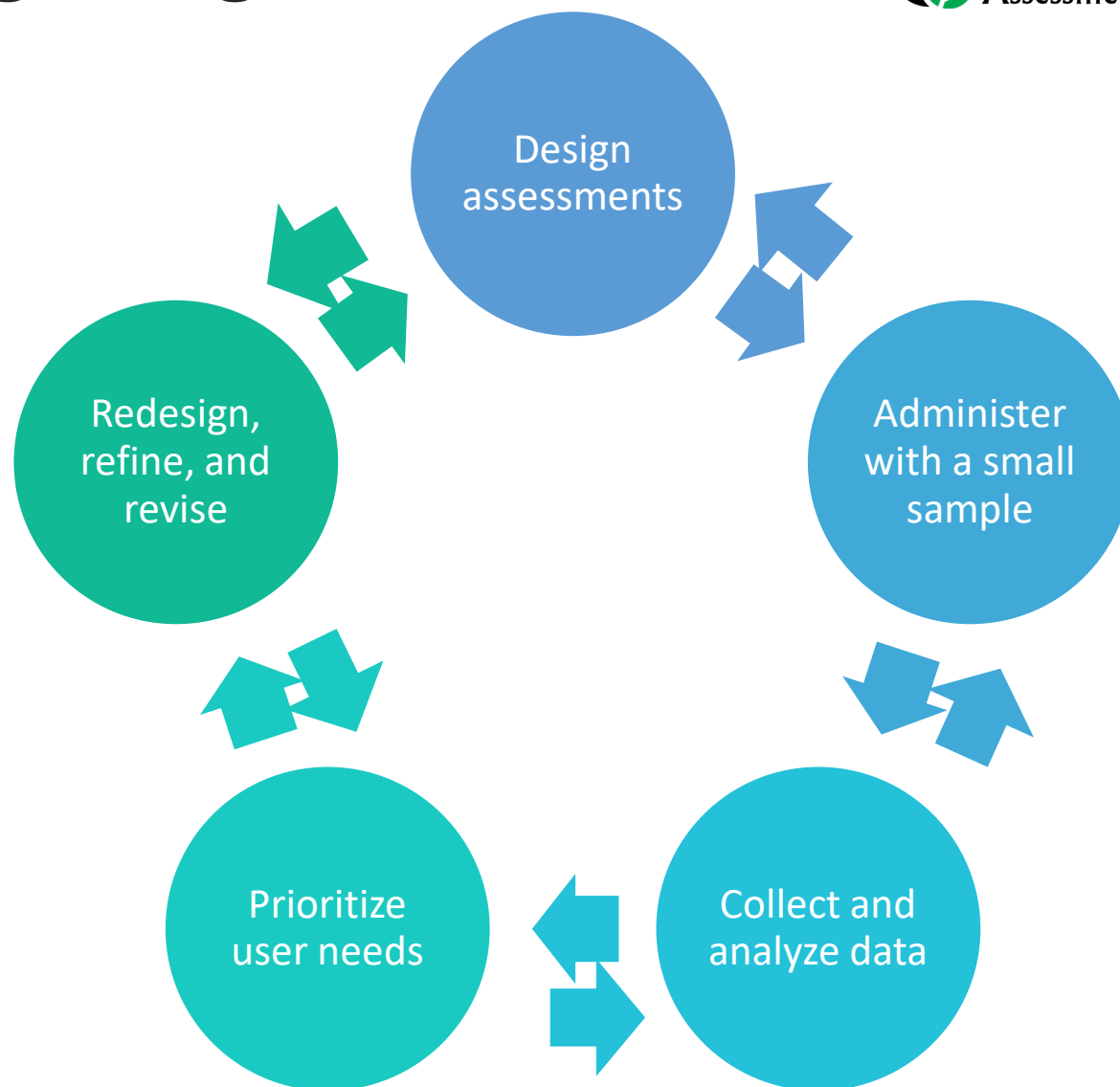
***Hypothesize  
Discover  
Plan***



# “Prototyping” Stage



***Test  
Adjust  
Create***





# Session 6 Focus

---

*"Design is not just what it looks like and feels like. Design is how it works."* -- Steve Jobs, co-founder of Apple, Inc.

## A Design Challenge

- Innovative Assessment Demonstration Authority (Every School Succeeds Act)
  - Background
  - Guardrails
  - Interview and Commentary
- Q&A / Closure

# Session 6 Focus

---

*"A designer knows he has achieved perfection not when there is nothing left to add, but when there is nothing left to take away."* -- Antoine de Saint-Exupéry, French writer and poet

## **The Outlook for Innovation – Challenges and Opportunities**

- Challenges and Inhibiting Factors
- Innovation in the Future
- Leveraging Lessons for Innovation
- Q&A / Closure

# Exploring the IADA Innovation Process – *Challenges and Opportunities*





# ***The Innovative Assessment Demonstration Authority Under the Every Student Succeeds Act***

Scott Marion & Carla Evans

National Center for the Improvement of Educational Assessment

*Reidy Interactive Lecture Series: Session 6*

September 24, 2021







## Session 6: Exploring the IADA Innovation Process – *Opportunities and Challenges.*

### Part 1: *Challenges*

Friday, September 24, 2021; 1:00-2:30 p.m.

Center Staff: Carla Evans & Scott Marion

The Innovative Assessment Demonstration Authority under the Every Student Succeeds Act was purportedly created to provide space for states to design and

insights regarding the potential for innovation.

#### Materials:

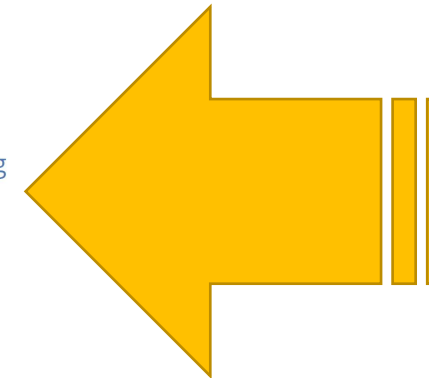
Brief Summary of IADA Programs

#### Videos:

Tensions Over Multiple, Intended Uses: Informing Instruction and Summative Accountability  
Scaling the Innovation Statewide  
Psychometric Scaling  
Comparability Challenges  
Advice to States  
Advice to the US Department of Education  
Concluding Thoughts: Tensions, Tradeoffs, and Innovation

## Session Materials:

<https://www.nciea.org/events/2021-rils-design-innovation-educational-assessment-systems>



## Referenced and Additional Resources

---

- U.S. Department of Education [IADA General Requirements & Recipients](#)
- Georgia's [Comparability Evidence Criteria & Evidence Document](#)
- Centerline Blogs about assessment innovation and IADA:
  - <https://www.nciea.org/blog/innovative-assessment/reality-faced-innovators-educational-assessments>
  - <https://www.nciea.org/blog/innovative-assessment/reality-innovation-educational-assessment>
  - <https://www.nciea.org/blog/essa/education-innovators-dilemma>
  - <https://www.nciea.org/blog/being-innovative-under-essas-innovative-assessment-demonstration-authority>
  - <https://www.nciea.org/blog/assessment-reform/accountability-roadblock-assessment-reform>
  - <https://www.nciea.org/blog/innovative-assessment/implementing-innovative-state-assessment-system-while-were-still>
- Centerline Blogs about tensions between instructional and summative accountability uses of assessment information:
  - <https://www.nciea.org/blog/state-testing/why-has-it-been-so-difficult-develop-viable-through-year-assessment>
  - <https://www.nciea.org/blog/state-testing/could-two-through-year-assessment-designs-provide-both-summative-and>



# Types of Assessment Flexibility under ESSA

---

Assessment flexibility offered under ESSA:

- ✓ High school assessment
- ✓ Use of interim assessments
- ✓ **Assessment innovation**
  - ✓ **Section 1204:** Innovative Assessment Demonstration Authority

# Innovative Assessment Demonstration Authority

---

- Allows for a pilot for **up to seven (7) states** (up to 4 can be part of a consortium) to use **competency-based or other innovative assessment approaches** for use in making accountability determinations
- Initial demonstration period of **three (3) years** with a two (2) year extension based on satisfactory report from the director of Institute for Education Sciences (IES)
- Subject to assessment **peer review** process, if it becomes the state assessment

# IADA Key Requirement & Flexibility

---

**Key requirement:** produce student-level annual determination of proficiency

- Careful consideration around psychometric scaling: how to combine multiple pieces of evidence, sometimes collected across the school year, into one determination—or choose not to combine

**Key flexibility:** assessments are not required to be the same statewide

- States can pilot the assessment system with a subset of districts before scaling the system statewide

# Four Major Guardrails

## Assessment Quality

- System comprised of high-quality assessments that support the calculation of **valid** and **reliable** annual determinations, as well as provide useful information to relevant stakeholders

## Comparability

- Produce student-level annual determinations that are **comparable** across assessment systems

## Scale Statewide

- Must have a logical plan to scale up the innovative assessment system statewide

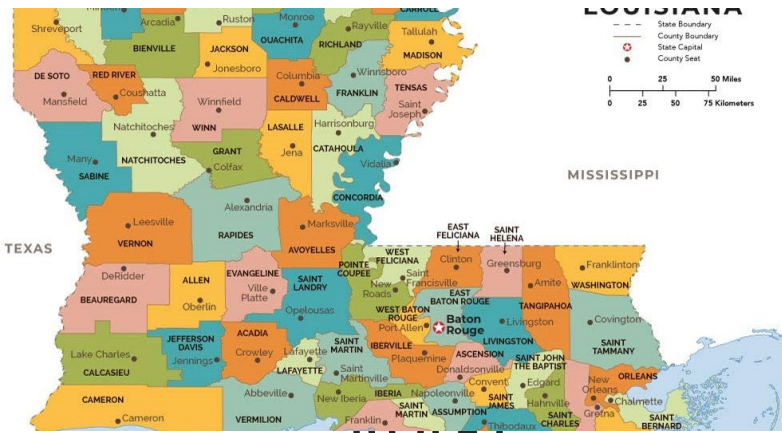
## Demographic Similarity

- Make progress toward achieving high-quality and consistent implementation across demographically diverse LEAs

# Current IADA State Recipients and Approaches

---





# IADA State Recipients



## Brief Summary of IADA Programs Approved (as of Sept 2021)

<https://oese.ed.gov/offices/office-of-formula-grants/school-support-and-accountability/iada/>

### Approved to start for the 2018-19 school year

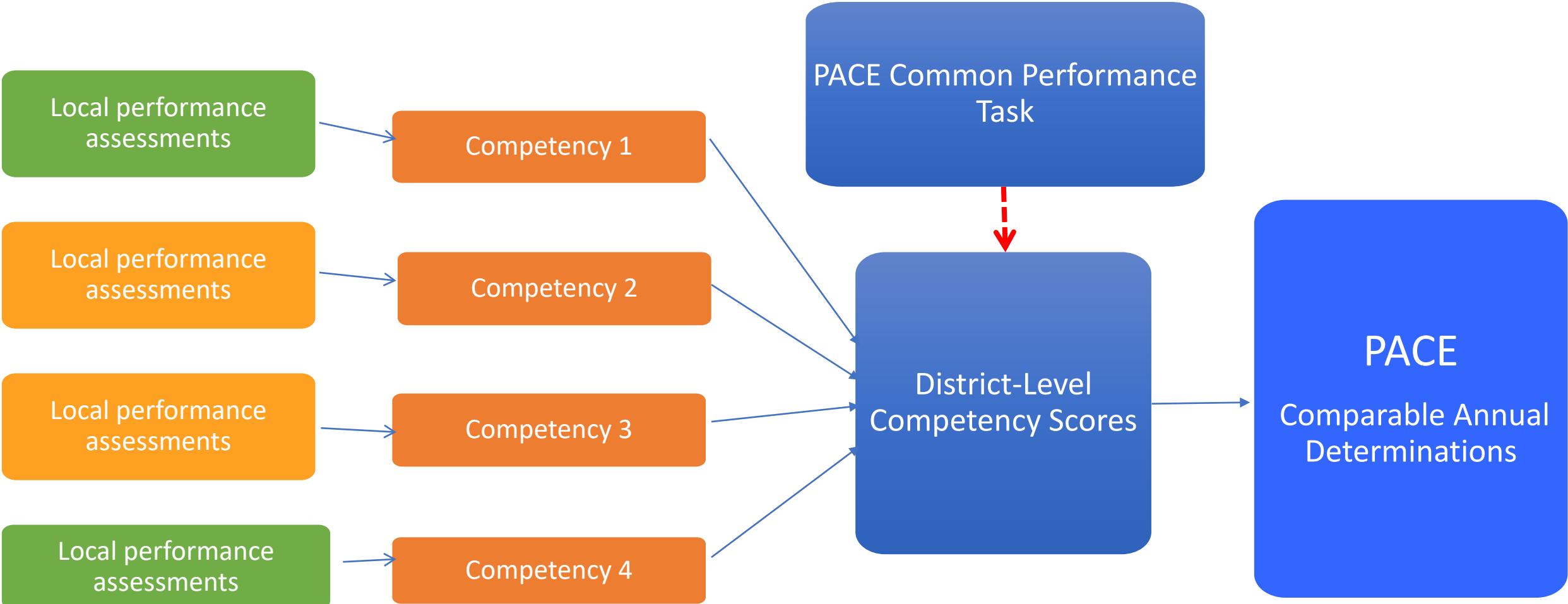
-**Louisiana**: through year model with three curriculum-embedded Gr 6-8 assessments in three testing windows over the year; eventually will scale to Gr 3-8 and include over end-of-unit assessment in fall and early winter; spring end-of-year assessment system is aligned with state-approved (Tier 1) and freely purchased (Tier 2) curriculum system is expanding to another curriculum) in ELA and math. This system is more equitable if all students have opportunity to come into an assessment system with the same knowledge. Hoping for full year of pilot data from 2021-22 SY for a psychometric scaling approach to produce annual determinations.

-**New Hampshire**: Performance Assessment of Competency Education (PACE) started in 2014-15 SY under NCLB; through year model in ELA, math and science in Gr 3-8; performance tasks created for high school to promote coherence, but no data submitted for accountability purposes; students take once per grade span state standardized test; in all other federally-required grades/subjects annual determinations based on classroom



# The New Hampshire PACE Assessment System

State summative assessment in select grades

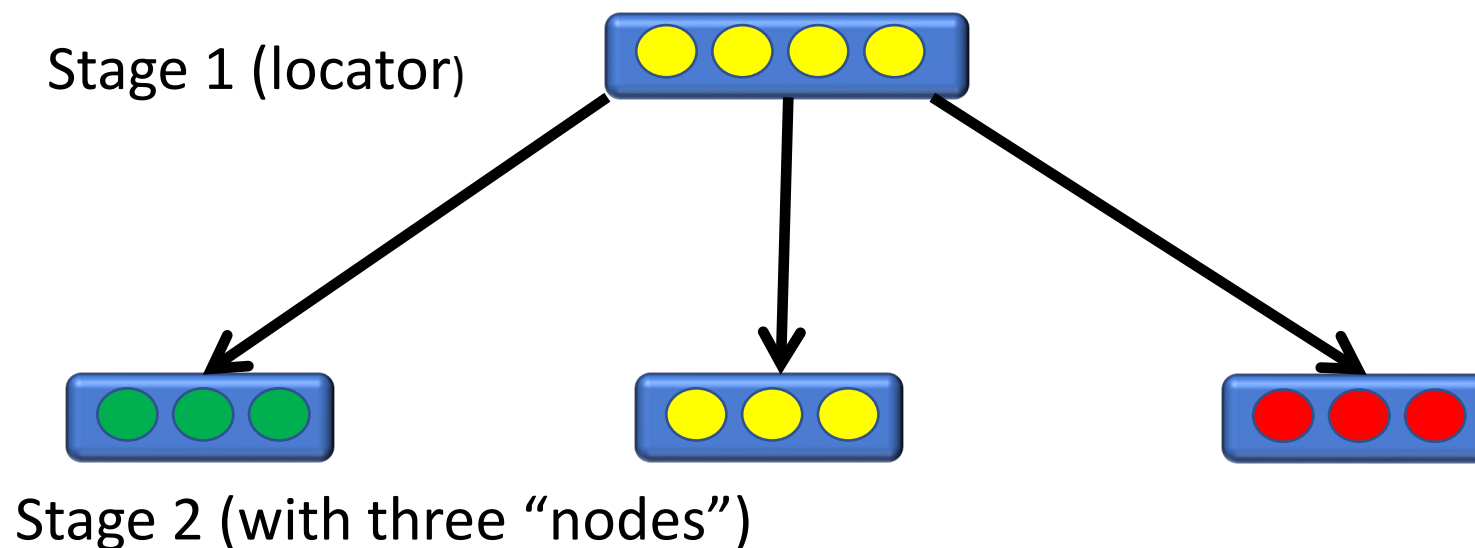


# Louisiana's Innovative Assessment Pilot



- Created and scaled quality, text- and content-rich curriculum (ELA Guidebooks) aligned to LA State Standards
- Built assessments with a focus on quality (LEAP 2025)
- Assessment system is comprised of end-of-unit assessment tied to units of instruction
- Supported implementation of quality curriculum with aligned PD

# North Carolina's "extended" multi-stage adaptive test



*Building on the foundation of NC Check-Ins, an assessment administered as an interim that provides formative student-level information, the North Carolina Department of Public Instruction is piloting an innovative assessment that will optimize the purposes of formative, interim, and summative into one assessment.*

NC has proposed using the results from the fall and winter interim tests as "stage 1."

# Georgia: Piloting two systems

---

- Georgia took an unusual approach to the IADA. They are running two pilots, even though the law is clear that by the end of the demonstration period, there can be only one state assessment system.
  - NAAVY—an approach using diagnostic classification modeling (DCM) to define “mastery” on assessments tied to each standard in math and ELA (20-30 standards per content area per grade)
  - GMAP—based on NWEA’s Measures of Academic Progress (MAP), this is somewhat like the NC model involving testing 3x/year, potentially using the multi-stage design to replace the single summative test.

# Massachusetts

---

- Focused on science
- Will maintain a shortened (½ length) MCAS throughout the pilot
- Developing innovative technology-enhanced tasks
- Engaging in rapid prototyping to refine the tasks
- Tied to the Kaleidoscope Collective for Learning
  - The Kaleidoscope Collective is a pilot program created in November 2019 to nurture [Deeper Learning](#) in Massachusetts public schools.
  - Kaleidoscope builds partnerships with educators and leaders through professional development, coaching, guidance, tools, and resources -- all built on a foundation of educational equity.



## IADA State Interviews

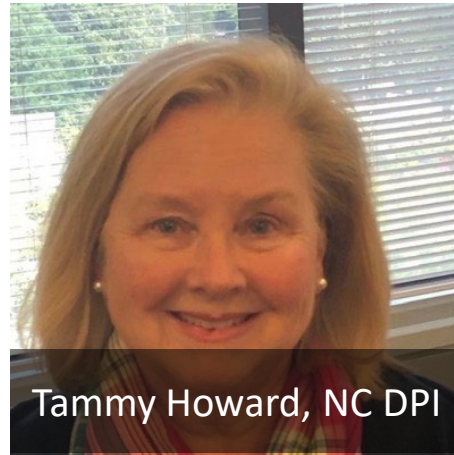
---

- We conducted interviews with four of the IADA approved states.
- We asked each state the same three questions:
  1. What have been the main challenges related to IADA implementation?
  2. How has your state planned to address challenges related to statewide scaling, psychometric scaling, and/or comparability?
  3. What advice would you give to other states interested in applying for the Demonstration Authority and/or the U.S. Department of Education?

## Seven Themes

---

- We categorized the responses into seven themes as depicted in the following video montages:
  1. Tensions Between Multiple, Intended Uses: Informing Instruction and Summative Accountability
  2. Scaling the Innovation Statewide
  3. Psychometric Scaling
  4. Comparability Challenges
  5. Advice to States
  6. Advice to U.S. Department of Education
  7. Concluding Thoughts: Tensions, Tradeoffs, and Innovation
- All videos can be accessed here: <https://www.nciea.org/events/2021-rils-design-innovation-educational-assessment-systems>
- We will take questions for us and/or the state leads at the end of all seven clips, so you can add them to the Q & A or just try to remember them



Tensions Over  
Multiple, Intended  
Uses: Informing  
Instruction and  
Summative  
Accountability



# Scaling the Innovation Statewide



# Psychometric Scaling





# Comparability Challenges



# Advice to States



# Advice to the U.S. Department of Education



# Concluding Thoughts: Tensions, Tradeoffs, and Innovation





Thank you to  
our IADA  
State  
Interviewees

#### Georgia

- **Allison Timberlake**, Deputy Superintendent for Assessment & Accountability, Georgia Department of Education
- **Garron Gianopulos**, Senior Psychometrician, NWEA
- **Michael Huneke**, Director of Assessment, Marietta City Schools, Georgia

#### Louisiana

- **Chanda Johnson**, Deputy Assistant Superintendent, Louisiana Department of Education

#### Massachusetts

- **Sam Ribnick**, Executive Director of the P20W Research and Data Hub, Massachusetts Department of Elementary and Secondary Education

#### North Carolina

- **Tammy Howard**, Director of Accountability, North Carolina Department of Public Instruction
- **Kinge Mbella**, Psychometrician, North Carolina Department of Public Instruction



# ***The Outlook for Innovation: Opportunities and Challenges on the Road Ahead***

Chris Domaleski

National Center for the Improvement of Educational Assessment

*Reidy Interactive Lecture Series: Session 6*

September 24, 2021





# Outline

---



Why does innovation fail?



Are there some potential areas  
of innovation on the horizon?



What are some strategies to  
promote success?

# Top seven reasons educational innovations fail...



1. Lofty goals but poorly specified plan



2.  
Insufficient  
stakeholder  
involvement  
and support



3.  
Incompatible  
with current  
conditions



#### 4. Lack of resources/ capacity





## 5. Constant change



6.  
Inattention  
to scaling



## 7. Insufficient monitoring and improvement

I STARTED THE DAY WITH  
LOTS OF PROBLEMS.  
BUT NOW, AFTER HOURS  
AND HOURS OF WORK,  
I HAVE LOTS OF PROBLEMS  
IN A *SPREADSHEET*.



What's the outlook for  
innovation in the  
future?



## What are the ‘wicked problems’ we need to solve?

---

- How can assessments better inform instructional interventions? (True ‘diagnostic’ assessment)
- How can we build and sustain capacity at scale to improve classroom assessment practices?
- How can we keep large scale summative assessment and high-stakes accountability from sucking all the oxygen out of the air?
- How can we improve monitoring and support, particularly to improve outcomes for traditionally underserved students?

### **WICKED PROBLEM.**

noun | wi-kəd | prā-bləm

a problem that is difficult or impossible to solve because of incomplete, contradictory, and changing requirements that are often difficult to recognize.

## Potential Assessment Innovations?

---

- **Utility:** Focus on providing more instructionally useful and timely information to inform instruction
  - Assessments supported by a learning progression with close links between curriculum and instruction
  - Systems to support personalized learning paths
- **Capacity:** Focus on investing in classroom assessment literacy skills
  - Innovations in training models (hint: it's not train the trainer)
  - Growth in customizable open access resources
- **Efficiency:** Focus on “right sizing” state summative testing
  - Grade-band testing
  - Matrix-sampling



# Potential Accountability Innovations?

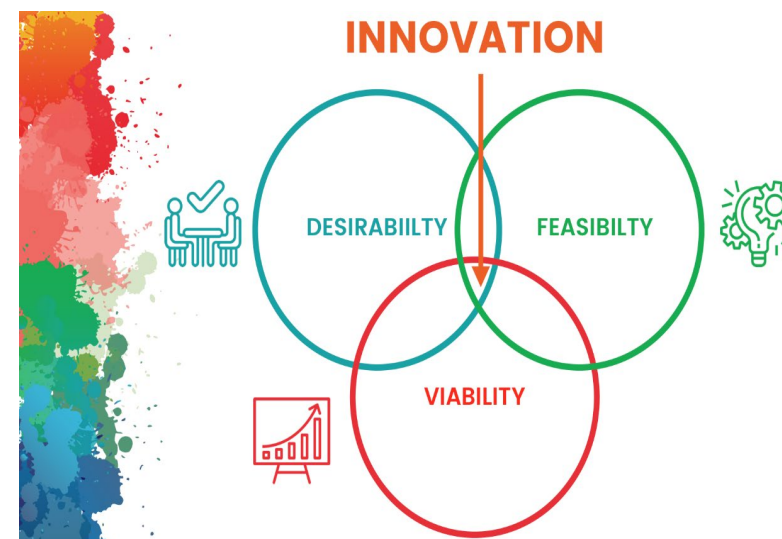
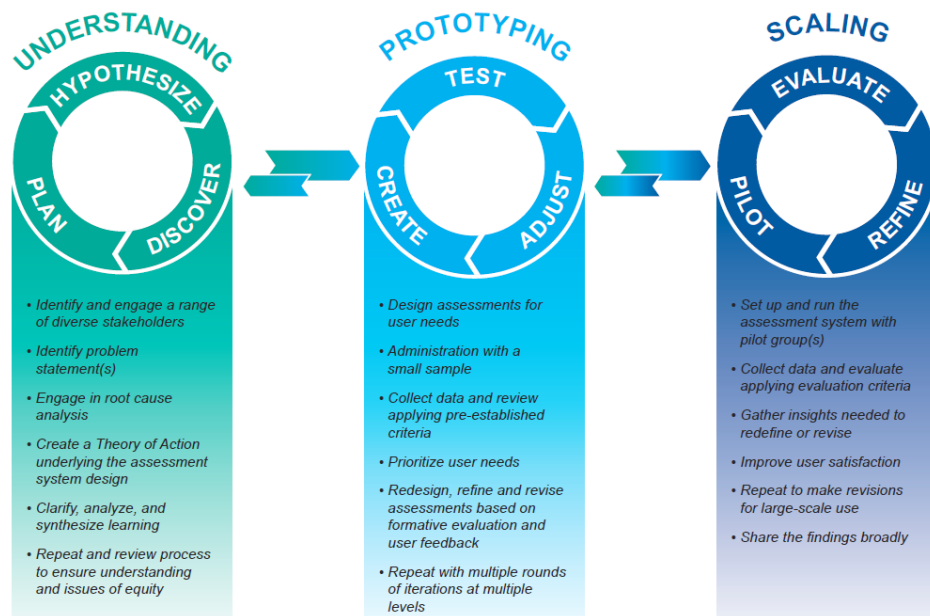
---

- **Coherence:** Focus on providing systems of accountability that are horizontally and vertically coherent
  - District and school systems, each customized to address the distinct mission and role
    - Stop treating districts like super schools
    - Recognize differences among schools (e.g. alternative)
  - Rethink the role of the SEA
- **Reciprocity :** Focus on specifying the conditions and supports required for success
  - Address inputs and outcomes
  - Build supports and interventions into the theory of action
- **Utility:** Focus on providing clear, actionable information and support to a range of stakeholders
  - Broaden information
  - Continued improvements in accessible and dynamic reporting systems
  - Investments in training and support

# Strategies to Promote Success

Let's return to some of the concepts Damian Betebenner and Jeri Thompson introduced in Session 1

## DESIGN INNOVATION PROCESS



## Final Thoughts

---

- *Think of innovation as a verb, not a noun.*
- *It's unlikely the process is linear, it's an unpredictable, twisty journey.*
- *First Do No Harm: Interactions and unintended consequences are always a threat*
- *Assessment (alone) is not an intervention. Connections to curriculum and instruction are critical.*

# Facilitator: Leslie Keng



# Closure – Session 6

---

## Takeaways :

- What are the design challenges under IADA?
- What are the opportunities for innovation in the near future for both assessment and accountability?



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