Assessing 21st Century Skills: Guidance for States

Dr. Carla Evans, Associate Director (cevans@nciea.org)

Center for Assessment

NASBE Virtual Meeting

September 25, 2025

National Center for the Improvement of Educational Assessment, Inc.

Established Expertise

A NH-based non-profit established in 1998. Improves assessment and accountability practices.

Extensive Reach

Long-term contracts with states (>35), districts, and other organizations to provide technical, design, and policy support.

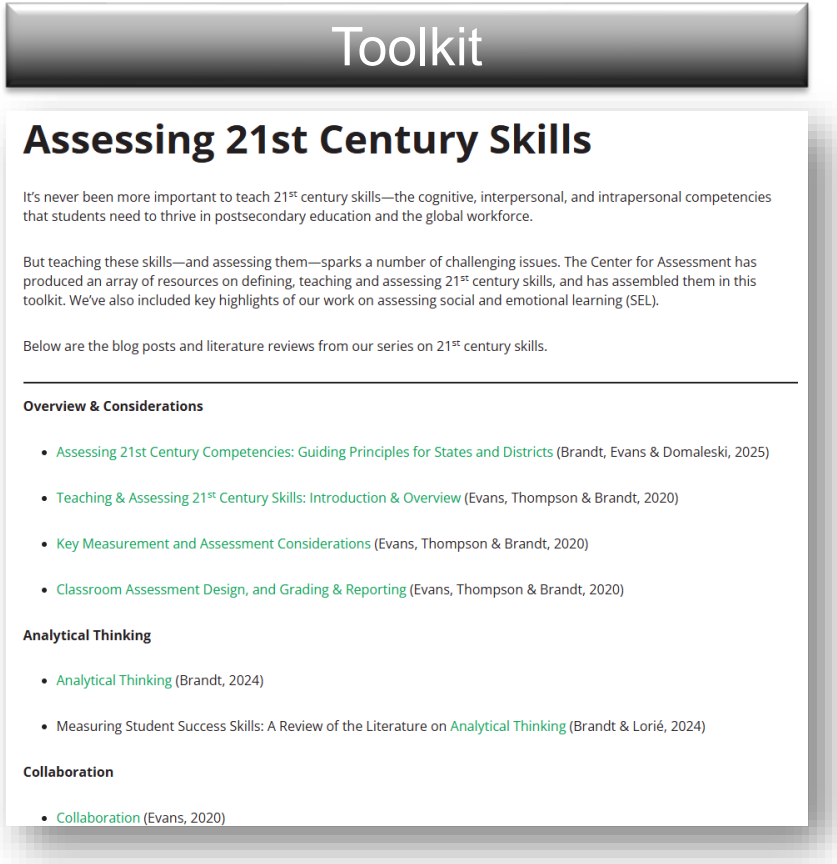
Purposefully Small

15 full-time professionals. All have advanced degrees and real-world experience.

Independent

Non-partisan and independent of governmental agencies or testing companies. All content is open source.

Assessing 21CCs Toolkit



In the toolkit you will find:

- State and District Guidance Paper
- Literature reviews on 21CCs
- Short blog posts summarizing key takeaways from the literature reviews

https://www.nciea.org/library/assessing-21st-century-skills/

ASSESSING 21ST CENTURY COMPETENCIES:

Guiding Principles for States and Districts

January 2025

Chris Brandt, Carla Evans & Chris Domaleski

National Center for the Improvement of Educational Assessment

Center for Assessment

National Center for the Improvement of Educational Assessment
Dover, New Hampshire

Center for Assessment

www.nciea.org

Roadmap

1. Assessment/Measurement Challenges
2. Recommended First Steps for States
3. Selected Examples
4. Concluding Thoughts

Assessment/Measurement Challenges

Basic Requirements for Sound Measurement

Clear construct definitions	*What is it you want to measure?*
Research-based grade-level, grade span, or end of HS proficiency targets	*What are students expected to know and do, and at what level of cognitive rigor at specific markers in time?*
Identified claims and uses	*What is it that you want to say about students, and for what purpose?*
Adequate domain mapping	*What collection of evidence is adequate to support identified claims and uses?*
Reliable scoring	*What degree of score reliability is adequate to support identified claims and uses?*

Measurement Challenges

Clear construct definitions	*Definitions vary across research traditions, domains, and cultures (leads to the jingle-jangle fallacy).*
Research-based grade-level, grade span, or end of HS proficiency targets	*Limited understanding about how these competencies develop and should be taught (learning progressions).*
Identified claims and uses	*Complex constructs are difficult to tease apart and accurately assess or measure in isolation. Many 21CCs are hard to observe because they represent internal processes that may not show up in tangible student work products.*
Adequate domain mapping	*Sufficient evidence to support a general claim ("This student is an effective collaborator") is difficult since 21CCs are inseparable from content and context, limiting the extent to which any single assessment can be used to make general ability claims. Could you collaborate as effectively if this room was filled with aeronautical engineers?*
Reliable scoring	*Application of knowledge and skills requires item types beyond multiple choice and short constructed responses. However, performance tasks, capstone projects, and portfolios are resource- and labor-intensive to design, implement, and score.*

Recommended First Steps for States

The question is not WHETHER a state or district CAN assess 21CCs.

The question is FOR WHAT PURPOSE, and what are the likely (un)intended consequences and tradeoffs?

This is similar to the discussions about the use of performance assessments in state testing.
See blog: "What Happens to Performance Assessment If We Use It for Accountability" (Evans, 2023).

Three Questions States Should Start to Ask & Answer

Center for Assessment

1. What is the problem we are trying to solve?

2. What are the constraints and requirements of an acceptable solution?

3. What would success look like?

1. What is the problem we are trying to solve?

- **Clarify Vision:** What is our state vision related to 21CCs? What are we hoping to accomplish and why? What is the problem we are trying to solve in the educational system? (e.g., PoGs+)
- **Clarify Assessment Goals, Purposes, and Intended Uses:** Where does assessment fit into that vision? What problem are we trying to solve with assessment, in particular? For example: *Our state needs a solution related to assessing 21CCs to…*
  - Get better information to support instruction and student self-reflection
  - Support deeper learning efforts
  - Evaluate program effectiveness
  - Hold schools accountable for incorporating 21CCs into all content areas
  - Other…

2. What are the constraints and requirements of an acceptable solution?

- **Requirements tell us what the system must do.**
  - What are the purposes the system MUST serve?
  - What are the required uses of the results?

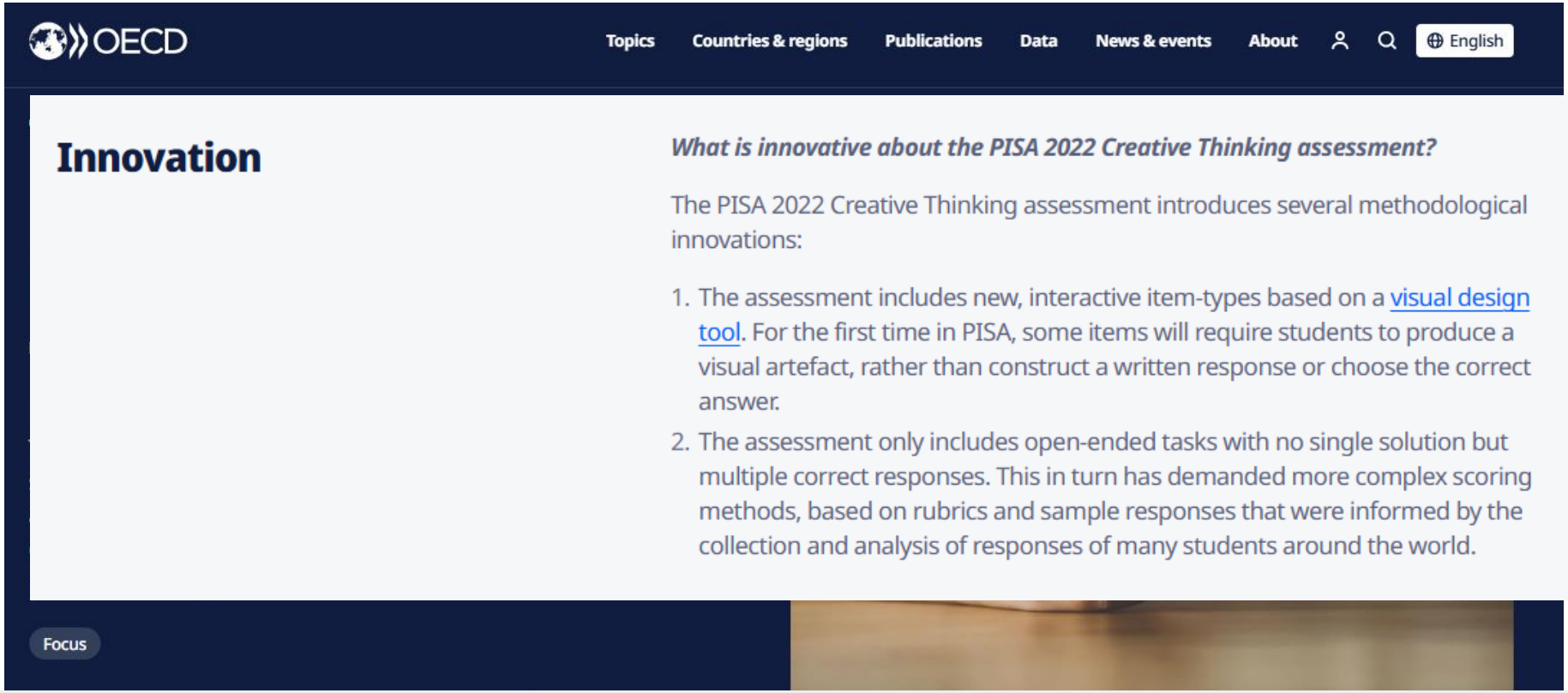
- **Constraints describe our limitations, such as:**
  - Total testing time (is there an upper limit)
  - Costs/Financial resources available
  - State/local capacity for local data collections and scoring efforts
  - Others?

Can we describe what success would look like?

- How would we know if we have succeeded?
- What are our success criteria?

Selected Examples

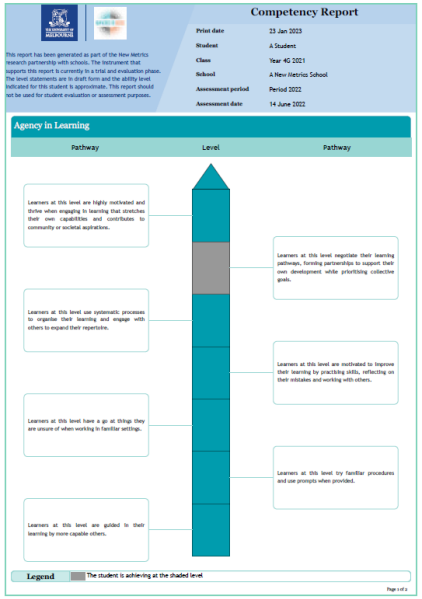
Large-Scale Quantitative Approach using Standardized Tests (e.g., OECD/PISA)

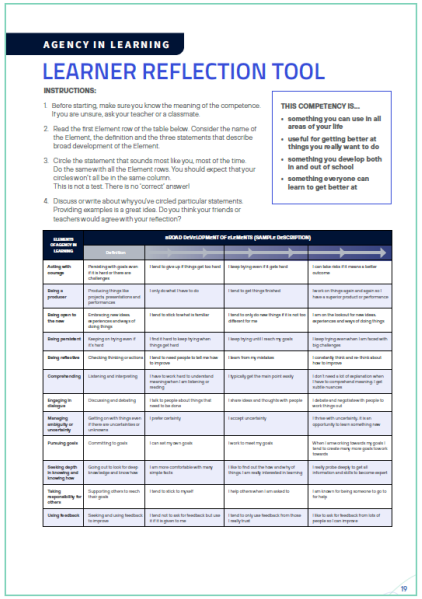


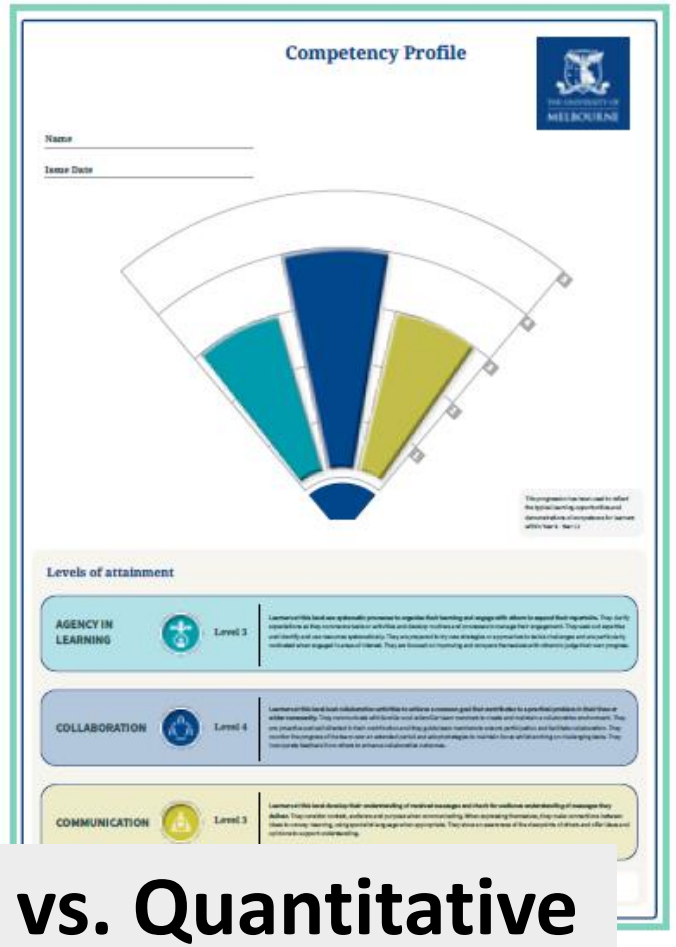
https://www.oecd.org/en/topics/sub-issues/creative-thinking/pisa-2022-creative-thinking.html

Specific vs. generalizable claims | Qualitative vs. Quantitative at the country level (15 years old)

Skills Profile Approach using Teacher & Student Self Report (e.g. Melbourne Metrics)







Specific vs. generalizable claims | Qualitative vs. Quantitative at the student level

Formative Feedback Approach using Student Work, Self Report, Observation, etc. against a Hypothetical Learning Progression /Continua (e.g., PBLWorks & IB)

Center for Assessment

Criteria	Less Sophisticated		Quality of Student Analytical Thinking			More Sophisticated
	1	2	3	4	5	6
	Select the column where most of the indicators describe what a student knows and can do independently					
Breaking a whole into parts	Correctly identifies the whole[3] without identifying any parts.	Identifies a limited number of parts (characteristics or functions) that make up the whole.	Identifies the important parts (characteristics or functions) of the whole.	Identifies the important parts (characteristics or functions) of the whole. Organizes parts into categories.	Identifies the important parts (characteristics or functions) of the whole. Organizes parts into relevant or meaningful categories that relate to one another and to the whole.	Identifies the important parts (characteristics or functions) of the whole and breaks them down further into relevant sub-parts. Organizes parts and sub-parts into relevant or meaningful categories or classifications that relate to one another and to the whole.
Examining the parts		Identifies a pattern or relationship between two or more parts.	Identifies most of the relevant patterns or relationships among the parts. Provides limited evidence to support conclusions	Identifies most of the relevant patterns or relationships among the parts. Provides sufficient evidence to support conclusions.	Identifies a comprehensive set of relevant patterns or relationships among the parts. Provides comprehensive and convincing evidence to support conclusions about how the parts relate to one another and the whole.	Identifies a comprehensive set of relevant patterns or relationships among the parts and subparts. Provides comprehensive and convincing evidence to support conclusions about how parts and subparts are distinct from one another, how they relate to one another, and how these distinctions/relationships work together to form (or support) a coherent whole.
Communicating		Explanation demonstrates a limited	Explanation demonstrates a	Explanation demonstrates	Explanation completely accounts for the important parts and how each part relates to the whole in	Explanation completely accounts for the important parts and subparts, their individual distinctions and relationships to other parts/subparts, and how they

Specific vs. generalizable claims | Qualitative vs. Quantitative at the student level

Selected Developmental Continua/Progressions Examples (Not Yet Validated)

Organization	Details
PBLWorks: https://www.pblworks.org/research/success-skills-rubrics	Critical thinking, collaboration, self-directed learning, complex communication, creativity Available by grade span: Grades K-2, 3-5, 6-12 and divided into 4 levels
ACARA (Australian Curriculum): https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/	Intercultural understanding, information and communication technology, critical and creative thinking, ethical understanding, personal and social capability, literacy, numeracy Divided into Levels 1-6 (end of foundation year to end of Year 10)
AAC&U VALUE Rubrics: https://www.aacu.org/initiatives/value-initiative/value-rubrics	Oral communication, problem solving, creative thinking, critical thinking, ethical reasoning, global learning, information literacy, inquiry and analysis, integrative learning, intercultural knowledge and competence, teamwork, written communication… [there are 16] Divided into 4 levels (rubrics are intended for postsecondary)
C!E: Essential Skills and Dispositions Developmental Frameworks: https://www.performanceassessmentresourcebank.org/system/files/EssentialSkillsandDevelopmentalFrameworks.pdf	Collaboration, communication, creativity, and self-direction Divided into 4 levels

Qualitative Descriptions Approach using Report Card Narrative

- A rich, contextualized, specific, and qualitative claim in a narrative report card statement.
- For example: “After receiving instruction on elements of effective collaboration in a project on transfer of energy, Susie demonstrated her knowledge and skill in helping the group to plan and make decisions; later reflecting on how she could better adjust her efforts next time to help the group accomplish its goal.”

Specific vs. generalizable claims | Qualitative vs. Quantitative at the student level

Competency-based Approach using Report Card Ratings

For Example:

Content area grades in the first part of the report card…

Somewhere at the end of the report card: 1 to 4 rating [or whatever rating scale] on…[insert the 21CCs you want]

- **Self-directed Learner:** The ability to be responsible for one's own learning.
- **Community Contributor:** The understanding that it is essential for human beings to work together.
- **Complex Thinker:** The ability to demonstrate critical thinking and problem solving.
- **Quality Producer:** The ability to recognize and produce quality performance and quality products.
- **Effective Communicator:** The ability to communicate effectively.
- **Effective and Ethical User of Technology:** The ability to use a variety of technologies effectively and ethically.

Specific vs. generalizable claims | Qualitative vs. Quantitative at the student level

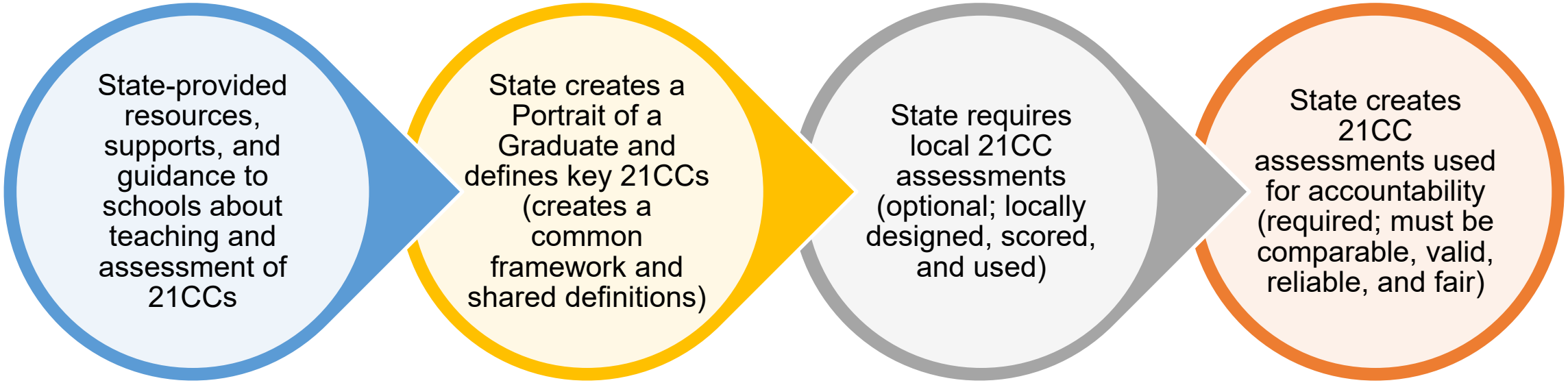
Other Ideas

- Portfolios of student work at the end of high school or other grade spans
  - See AP Studio Art – clear criteria for inclusion of student work; clear scoring criteria; rater calibration and training for reliable scoring.
- Graduation capstone projects at the end of high school
- Mix and match some of these ideas

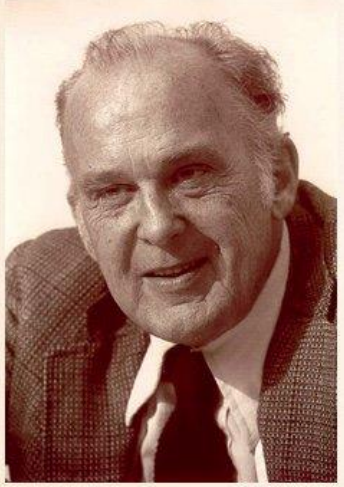
Concluding Thoughts

State Involvement Along a Continuum

States have different options to signal their policy values related to 21CCs. Some involve state-level assessment, while others do not. These are just some examples along a continuum:



Campbell's Law



Donald Campbell

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.

No Matter What, Piloting is Key

- It is always essential to pilot.
- Piloting the assessment itself to collect validity evidence.
- If the assessment is used in an accountability system, piloting how to include results, analyze impact on school ratings, and gather technical quality evidence.

Thank you!

Carla Evans (cevans@nciea.org)

www.nciea.org