



Challenges and Opportunities in Next Generation Assessment Systems: the Consortia and Beyond

Pascal D. Forgione, Jr., Ph.D.
Distinguished Presidential Scholar and Executive Director
The K-12 Center at ETS
Austin, Texas

TSNAP Luncheon, Keynote Presentation
Austin, Texas
February 17, 2015

Driving Advances in K-12 Assessment



Presentation Outline

- Key Drivers of Next-Gen K-12 Assessments
- The “Call to Action”
- Next Generation Assessment Systems
- Comparison of Traditional and Next-Gen State Assessment Items
- Looking to the Future



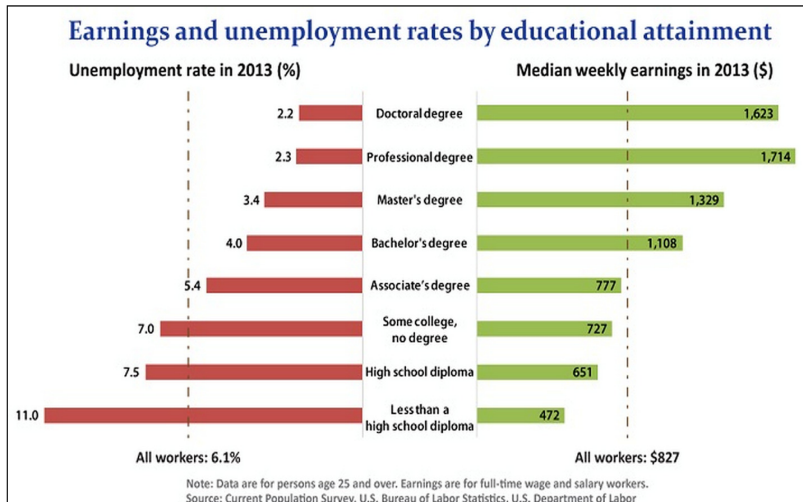
Presentation: Part 1

- **Key Drivers of Next-Gen Assessments**
 - Workforce demands
 - Post-secondary remediation rates
 - State academic performance expectations
 - High school – post-secondary skills gap
- The “Call to Action”
- Next Generation Systems of Assessments
- Comparison of traditional and next-gen state assessment items
- Looking to the Future

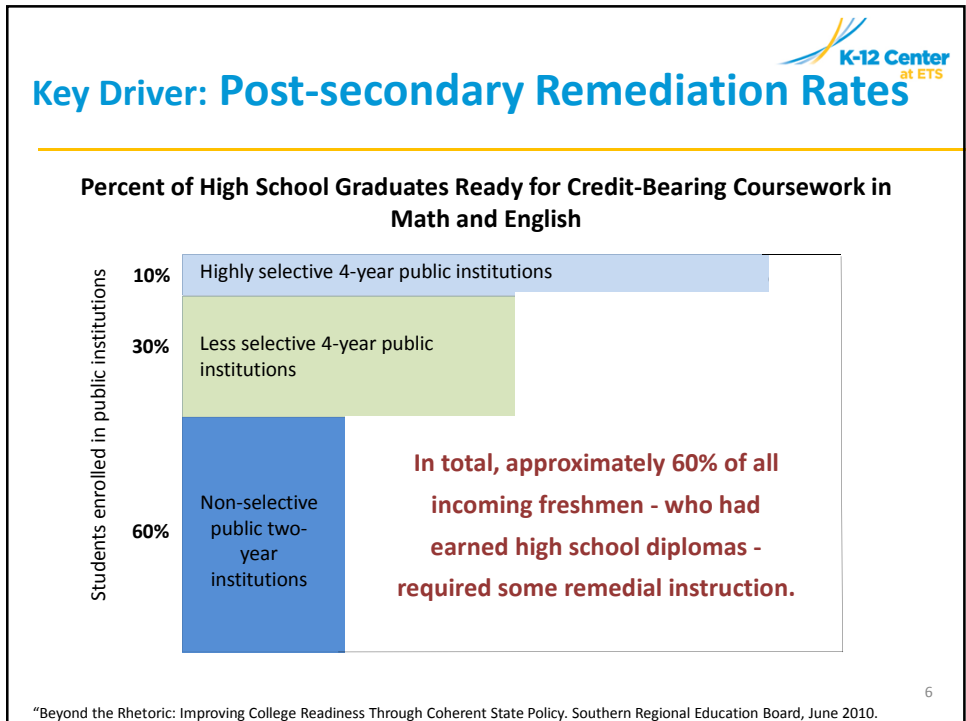
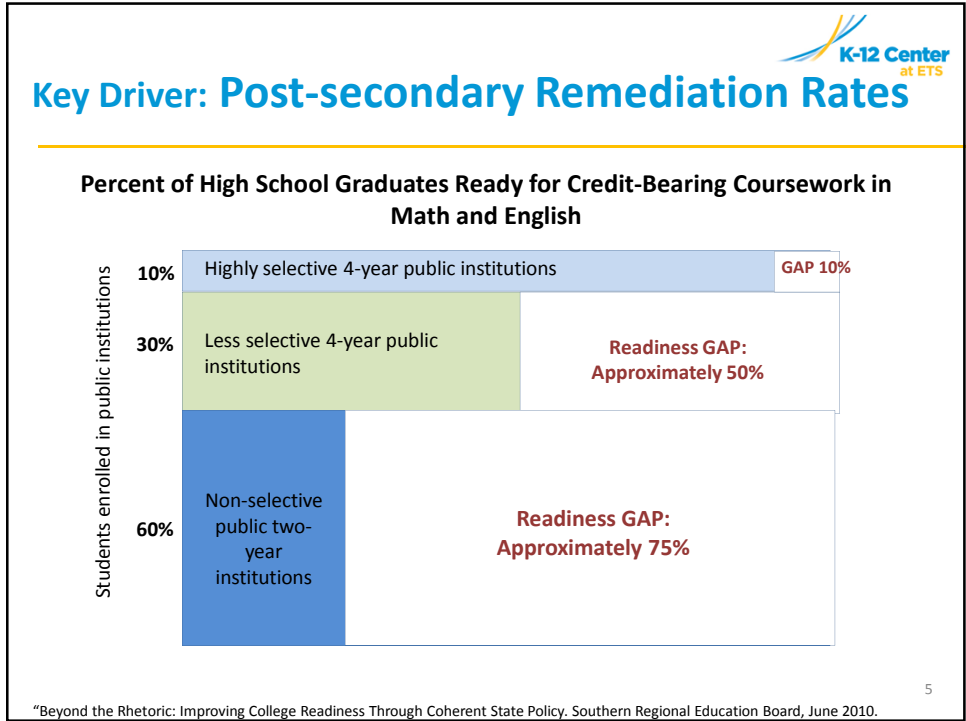
3



Key Driver: Workforce Demands



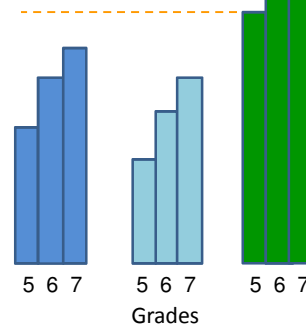
4



Key Driver: State Academic Performance Expectations Vary Widely



- Each state developed its own content and performance standards and contracted for its own assessments*
- Huge variability found in the rigor of states' performance standards – as much as 3 to 4 grade levels in some cases†



* Exception: Several new England states developed common assessments.

† Gary W. Phillips, "International Benchmarking: State and National Education Performance Standards." American Institutes of Research, 2014.

7

Key Driver: High School - Postsecondary Skills Gap



What are the critical skills our high school graduates need?

Meta-analyses of college and workforce skills, knowledge and abilities (SKAs) point to a critical set of higher order skills:

- Problem solving (non-routine)
- Critical thinking
- Systems thinking
- Information/ ICT literacy
- Creativity/Innovation
- Learning to learn, meta-cognitive skills

Measurement of 21st Century Skills within the Common Core State Standards, P. Kyllonen, 2012

8

Key Driver: High School - Postsecondary Skills Gap



The identified “critical skills” tend to be Depth of Knowledge Level 3 or 4 skills.

Skills, Knowledge and Abilities (SKAs)	Depth of Knowledge (DOK)
Extended Thinking: Application, research, modeling, and problem solving skills	DOK 4
Strategic Thinking: Assess, compare, evaluate, hypothesize	DOK 3
Conceptual and Procedural knowledge: Classify, organize, summarize, infer	DOK 2
Basic skills: memorize, list, identify, tell	DOK 1

Webb, N.L. 2002

9

Key Driver: High School - Postsecondary Skills Gap



Have existing state tests measured these SKAs well?

- RAND 2012 analysis* of assessments from 17 states reputed to have higher standards and more ambitious assessments:
 - Only 21% of ELA items and
 - Fewer than 2% of mathematics items
 required such higher order skills (DOK 3 or 4).
- National Center for Research on Evaluation, Standards and Student Testing (CRESST)**:

“This is NOT sufficient to support college and career readiness.”

* Estimating the percentage of students who were tested on cognitively demanding items through the state achievement tests. RAND Corporation, 2012

** Research Perspectives on Next Generations Assessments, J.L. Herman, 2013

10



Key Drivers

Questions & Discussion

11

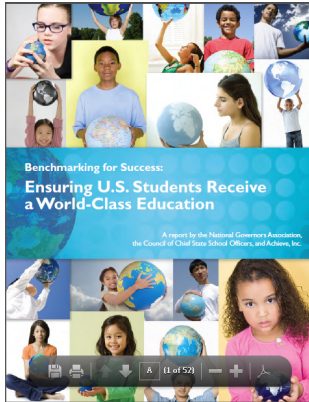


Presentation: Part 2

- Key Drivers of Next-Gen Assessments
- **The “Call to Action”**
 - More Rigorous Standards and Higher Quality Assessment
- Next Generation Systems of Assessments
- Comparison of traditional and next-gen state assessment items
- Looking to the Future

12

2008: The Call to Action by the NGA, CCSSO and Achieve



Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education. 2008

Called for:

- States to collaboratively develop and adopt a **common core of standards** in ELA and math that are aligned to college and workplace expectations
- The USED to provide funding for the **development of new assessments** that measure these standards
- States to leverage their collective influence and buying power to ensure high quality instructional and professional development materials

13

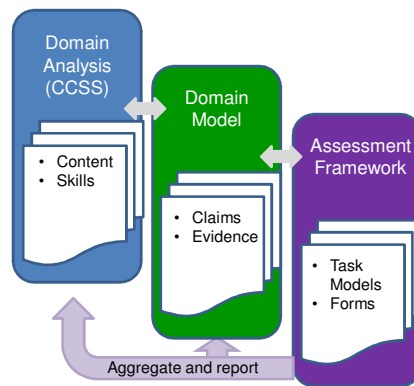
The Call to Action: Why New Assessments?



Old item development process: 1-to-1

- Items written to address individual Standards
- No well-defined process to develop complex tasks that assess multiple Standards
- No well-defined means by which the development process ensured higher level report claims were justified

Evidence-centered design*: more complex tasks and claims



*R. Mislevy et al, 1997
http://www.education.umd.edu/EDMS/mislevy/papers/ECD_overview.html

Attribution: Content of this slide adapted from WestEd presentation, 11.18.14

14



The Call to Action: Why New Assessments?

Will the Consortia assessments do a better job of assessing these critical higher order skills?

- CRESST 2013 review of Consortia work to date*:
 - 68% of ELA/literacy targets and *(RAND: 21% now)*
 - 70% of mathematics targets *(RAND: 2% now)*
 will be at higher order skill levels, DOK 3 or 4
- Independent reviews being conducted by Fordham Institute and HumRRO of PARCC, Smarter Balanced and ACT Aspire at grades 5, 8, and high school. Reports due summer 2015.
- CRESST conducting independent review of PARCC and Smarter Balanced assessment of higher order skills (report date TBD)

15

* On the road to assessing deeper learning: The status of Smarter Balanced and PARCC assessment consortia. Herman, J.L. and R.L. Linn, 2013



Call to Action

Questions & Discussion

16

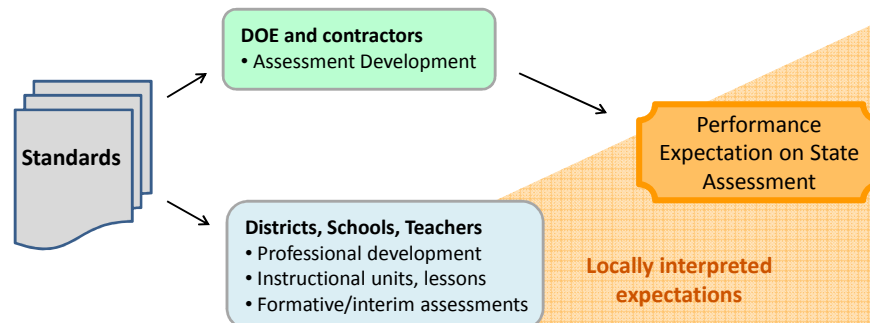
Presentation: Part 3

- Key Drivers of Next-Gen Assessments
- The “Call to Action”
- **Next Generation Systems of Assessments**
 - **Problems and Solutions of Next-Gen Assessments**
 - **The Federal Race To The Top (RTTT) Assessment Program**
 - **PARCC Design**
 - **Smarter Balanced Design**
 - **Key Similarities and Differences**
- Comparison of traditional and next-gen state assessment items
- Looking to the Future


17

Next Generation Systems of Assessments

**A significant problem to address:
“interpretation slippage”**

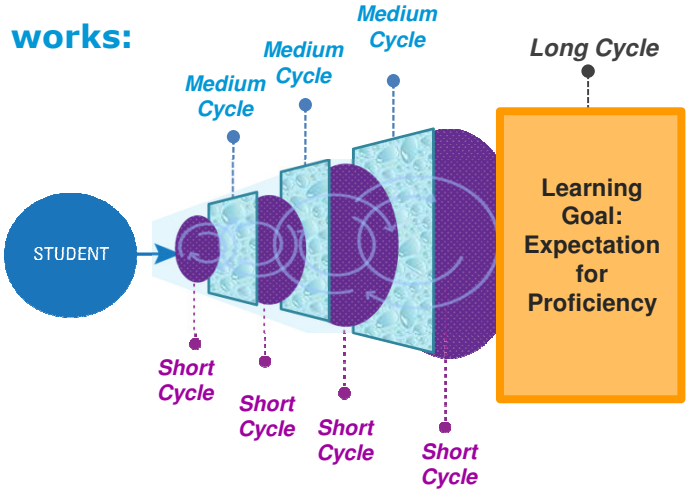


Based in part on A. Porter et al, *Alignment as a Teacher Variable*, 2007.




Next Generation Systems of Assessments

What works:



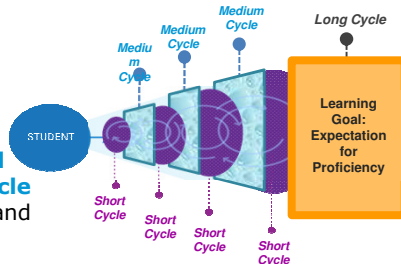
19

Formative Assessment for next Generation Science Standards: A Proposed Model. Presentation by J. Herman, 2013



Next Generation Systems of Assessments

- **Learning goal(s) and performance expectations are clearly articulated and communicated to teachers and students** (FAST, 2008)
- **Formative assessments are a coordinated continuum of short, medium, and long cycle activities** designed to both support learning and elicit evidence of it (William & Thompson, 2007)
- **A robust cognitive model** has been developed of how learning is expected to develop toward the learning goal(s) and **serves as the foundation for formative and interim assessment interpretation and feedback** (Herman, 2013)
- An **interpretive framework is provided and applied** to student responses to determine **where the student is relative to the goal and what is needed next** to move forward (Herman, 2013)
- Assessment activities **foster student motivation and engagement** (Harlen, 2006; Dweck, 2006)



20

Requirements of the Race to the Top Assessment Program



Announced in March, 2010, groups of 15 or more states could apply for a grant to develop online, next-generation assessment systems that:

- Assess **shared standards** in mathematics and ELA/literacy for college- and career-readiness
- Measure **individual growth as well as proficiency;**
- **Utilize technology** to the maximum extent appropriate; and
- Provide **information that is useful** in informing:
 - Teaching, learning, and program improvement;
 - Determinations of school effectiveness and of principal and teacher effectiveness for use in evaluations and support; and
 - Determinations of individual student college and career readiness, such as determinations made for high school exit decisions, college course placement to credit-bearing classes, or college entrance.
- **Operational by 2014 – 2015 school year**

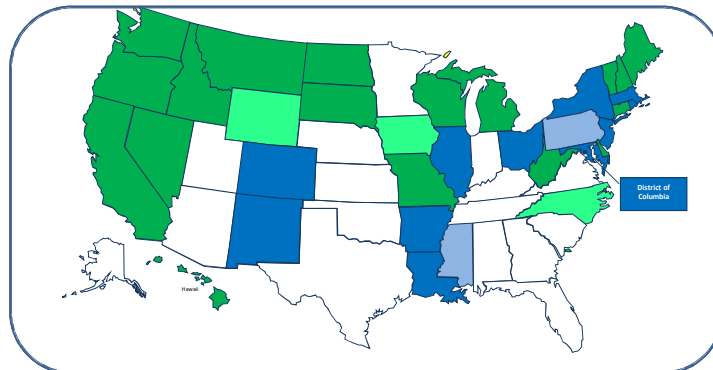
(US Department of Education, 2010)
21

Membership in Comprehensive Assessment Consortia



Membership as of February 5, 2015

- | | | |
|---|---|---|
| <p>■ PARCC
Governing: 11 states & DC*
Participating: 2 states*</p> | <p>■ Smarter Balanced
Governing: 18 states and USVI
Affiliate: 3 states</p> | <p> Neither:
16 states</p> |
|---|---|---|



* Pennsylvania have not been active of late.

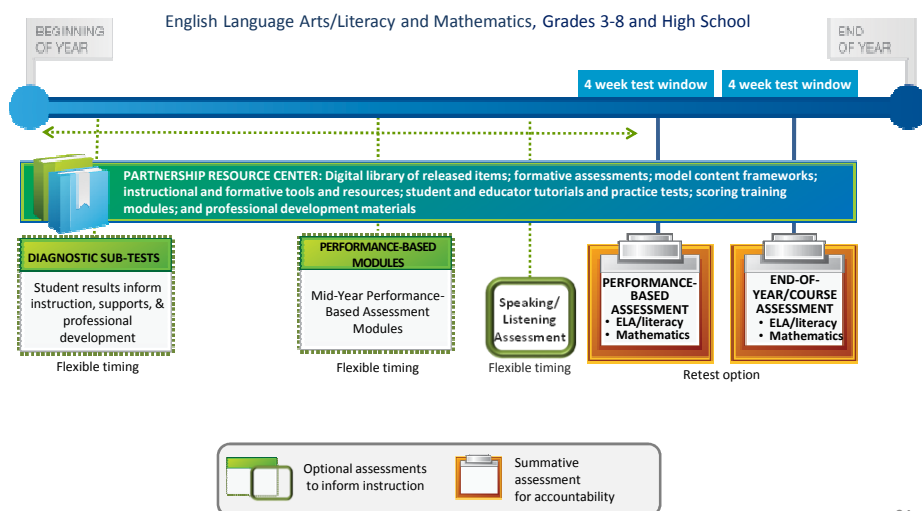
22

The Partnership for Assessment of Readiness for College and Careers (PARCC)

Details confirmed as of January 15, 2015

23

PARCC Assessment System

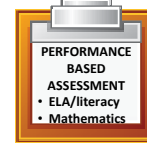


24

PARCC Summative Assessments: Performance-Based Assessments



- Taken after at least **75%** of instructional year
- **3 ELA/literacy tasks over 3 sessions**
 - 1 narrative writing task
 - 1 literary analysis task
 - 1 research simulation task
- **2 math test sessions with multiple tasks**
 - complex, real-world application problems
 - emphasis on mathematical practices and modeling
- **Electronic and distributed human scoring**
- **Results by end of school year**

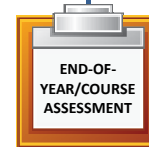


25

PARCC Summative Assessments: End-of-Year Assessment



- Taken between **75%** and **90%** of instructional year
- **Multiple item types** including constructed response and technology enhanced
- **Fixed-form**, with multiple forms per grade level/course
- Will include items that measure the **full performance continuum**
- **Electronically scored**
- High school: Also, available at end of 1st semester to support **block scheduling**



26

PARCC: Summative Components

- Given over nine testing sessions
- **Estimated testing time, ELA + math:** is expected to be 6.5 – 7.5 hours, although the scheduled time will be longer to allow for students who want more time
- **PBA and EOY scores combined for accountability**
- **1 Retest in 3 – 8 and up to 3 in HS,** as State/locally approved
- **Paper and pencil version** as an accommodation and for schools with State DOE approval

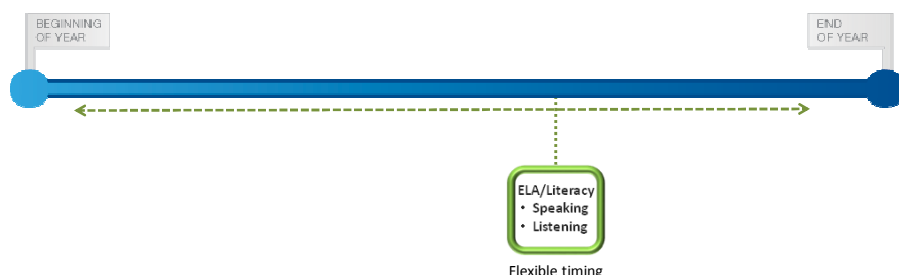
27

PARCC Supports: Interim Assessments

- **Optional Diagnostic Sub-Tests,** grades 2 – 8 (Fall 2015)
 - Two-stage adaptive tests designed to inform instruction
 - Sub-tests are: reading comprehension, reading fluency, decoding, vocabulary, mathematics comprehension, mathematics fluency
- **Optional Performance-Based Modules,** grades 3 – 11
 - Designed to mirror summative Performance Tasks
 - Scored by teachers

28

PARCC: Optional Speaking/Listening Assessment

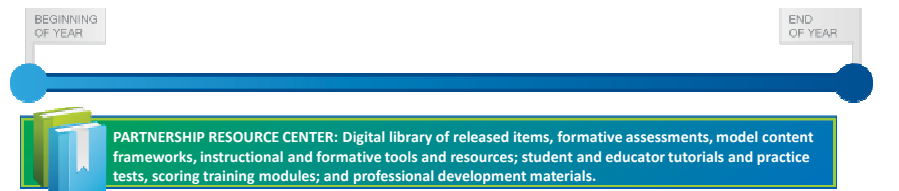


The diagram shows a horizontal timeline from 'BEGINNING OF YEAR' to 'END OF YEAR'. A green dashed double-headed arrow spans the entire duration. A green box labeled 'Flexible timing' is positioned below the timeline, with a dotted line connecting it to the center of the timeline. Inside this box, it lists 'ELA/Literacy' with sub-points for 'Speaking' and 'Listening'.

- Available for grades K-12 as an optional assessment -- **not used for accountability**
- **Scored by classroom teacher** using standardized rubric
- Scores may be used within students' grades

29

PARCC Supports: The Partnership Resource Center



The diagram shows a horizontal timeline from 'BEGINNING OF YEAR' to 'END OF YEAR'. A blue box with a folder icon on the left is positioned below the timeline. The text inside the box describes the Partnership Resource Center as a digital library of released items, formative assessments, model content frameworks, instructional and formative tools and resources; student and educator tutorials and practice tests, scoring training modules; and professional development materials.

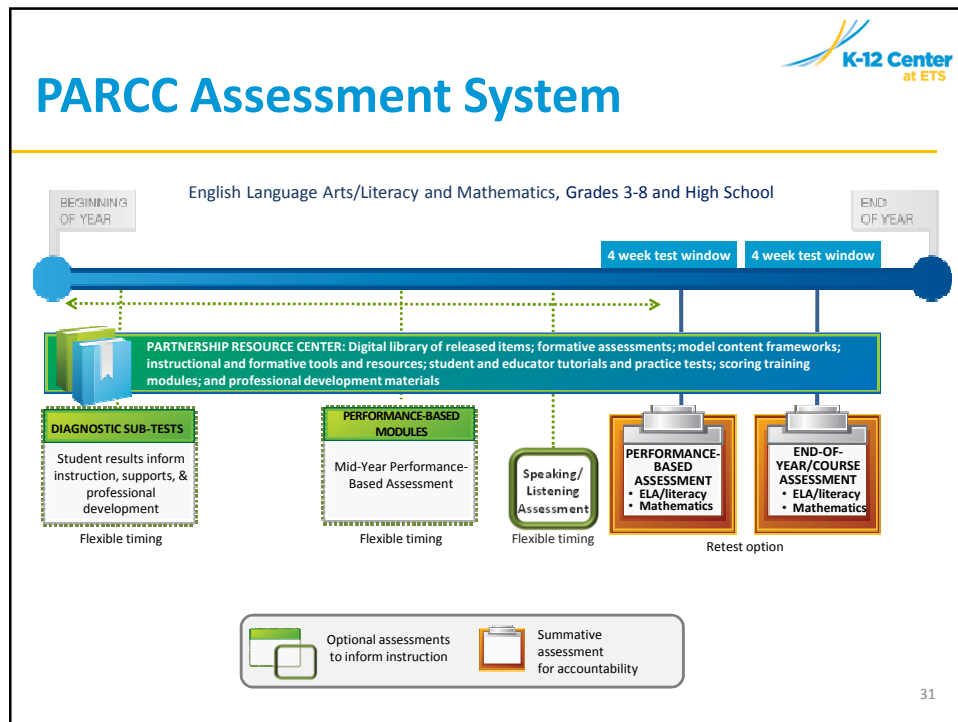
PARTNERSHIP RESOURCE CENTER: Digital library of released items, formative assessments, model content frameworks, instructional and formative tools and resources; student and educator tutorials and practice tests, scoring training modules; and professional development materials.

To be available summer 2015* and include:

- **Online practice tests for each grade/course**
- **Formative assessment items and tasks**
- **Professional development materials** – scoring, use of data
- **Instructional tools and resources** developed by Partner states
- Optional “ready-to-use,” instructionally embedded **formative tasks and tools for K-1** (Fall 2015)

* Some of these resources are available now at www.parcconline.org

30



PARCC: Supports and Timeline

Plans as of January 15, 2015, subject to change

2014 – 2015 school year

- **Mid-Year and Summative assessments**
- **Partnership Resource Center** launches summer 2015
- **College readiness tools** released (TBD)

2015 – 2016 school year

- **K-1 formative tools**
- **Optional Diagnostic Sub-tests and Speaking/Listening assessments available**

32

PARCC Assessment System



Questions & Discussion

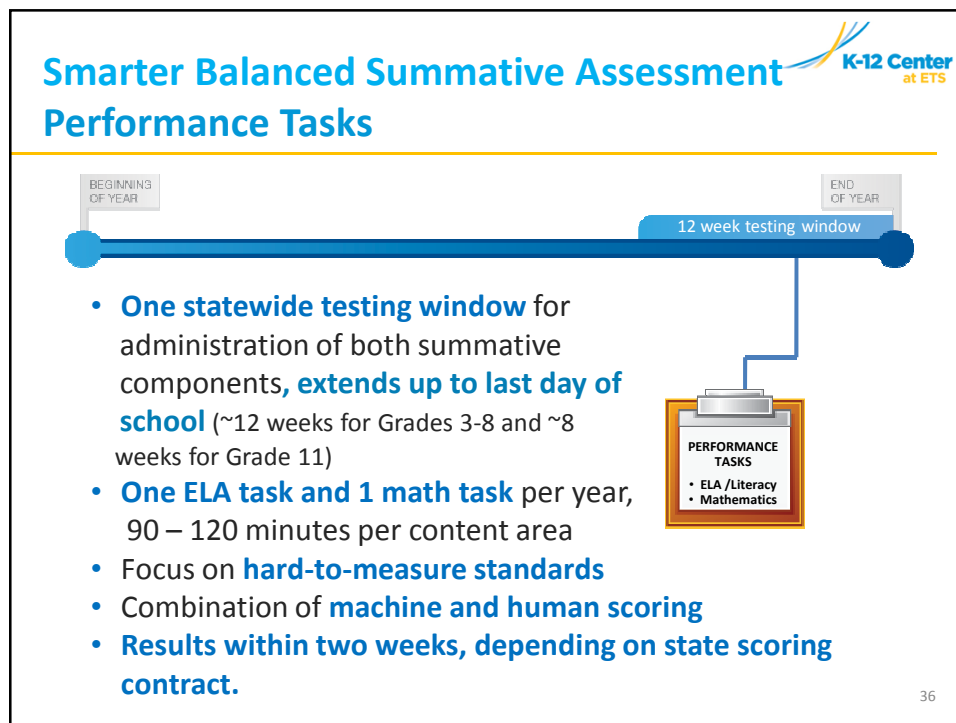
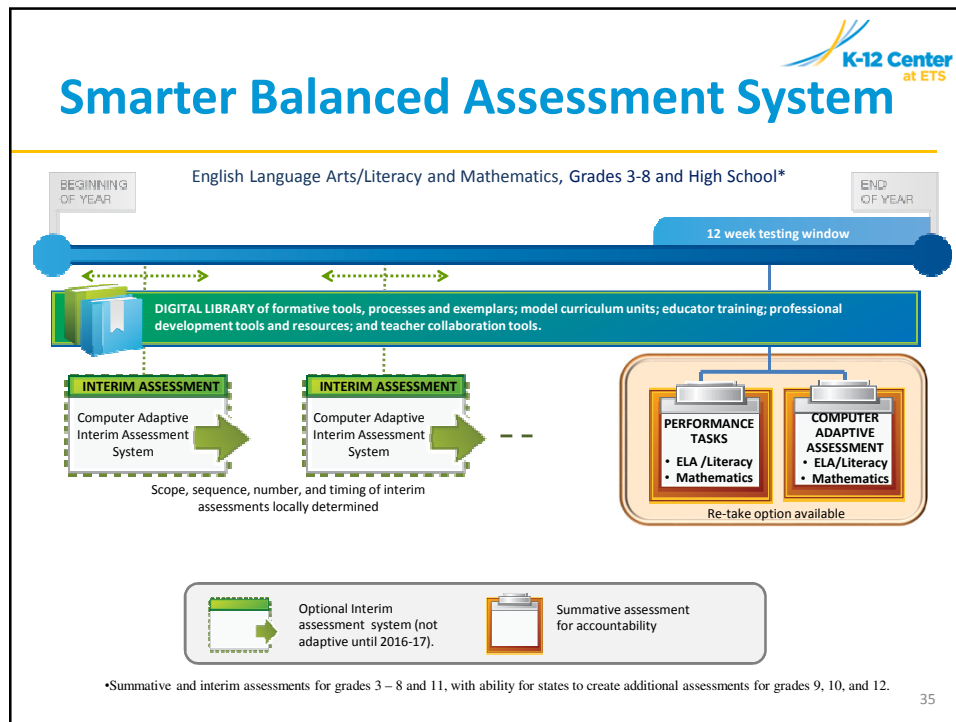
33

The Smarter Balanced Assessment Consortium

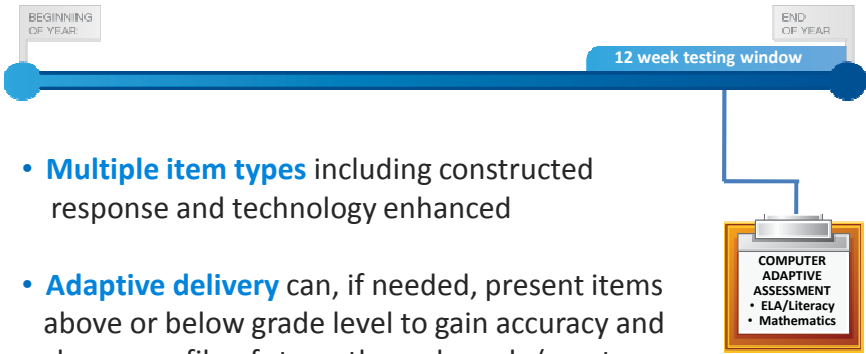


Details confirmed as of January 15, 2015

34



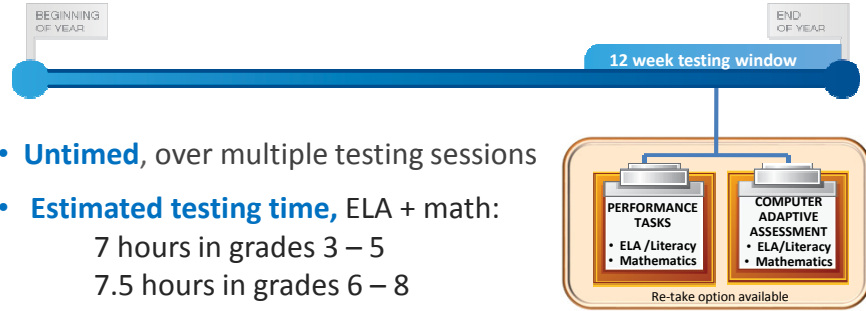
Smarter Balanced Summative Assessment Computer Adaptive End-of-Year Assessment



- **Multiple item types** including constructed response and technology enhanced
- **Adaptive delivery** can, if needed, present items above or below grade level to gain accuracy and clearer profile of strengths and needs (*most students will see only grade-level items*)

37

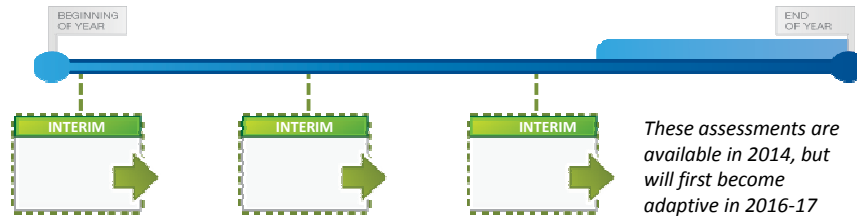
Smarter Balanced Summative Components



- **Untimed**, over multiple testing sessions
- **Estimated testing time**, ELA + math:
 - 7 hours in grades 3 – 5
 - 7.5 hours in grades 6 – 8
 - 8.5 hours in grade 11
- **PT and Computer Adaptive scores combined** for accountability
- **1 Retake**, if locally approved, for testing irregularities
- **Paper and pencil version for 3 years**; thereafter as accommodation

38

Smarter Balanced Supports: Optional Interim Assessment System

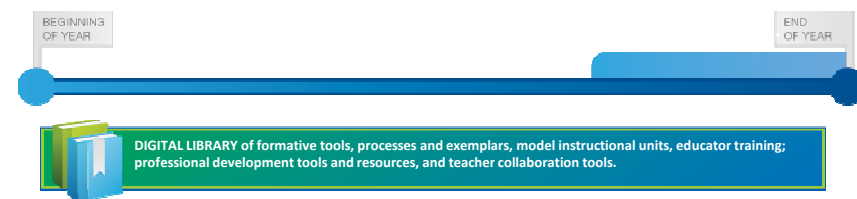


- **Two Options:**

- **Interim Comprehensive Assessments** use the same design as the summative assessments, assess the same range of standards, and provide scores on the same scale.
- **Interim Assessment Blocks** focus on smaller sets of related concepts and provide more detailed information for instructional purposes. There are between five and seventeen blocks per subject per grade
- Test questions are **not secure**, and there are no restrictions on the number of times that teachers may access assessments

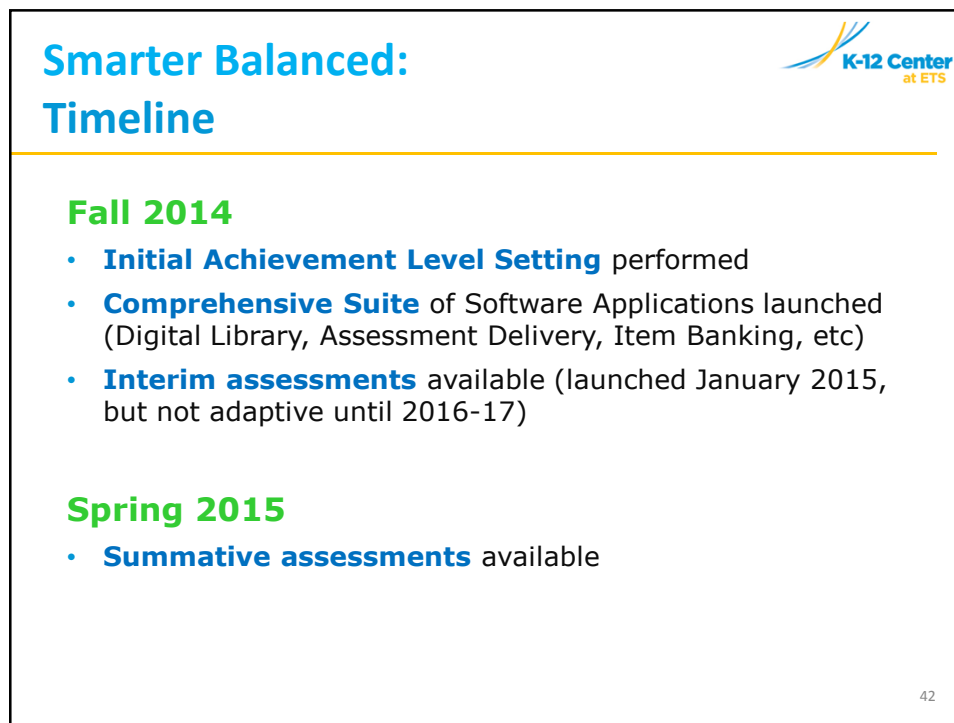
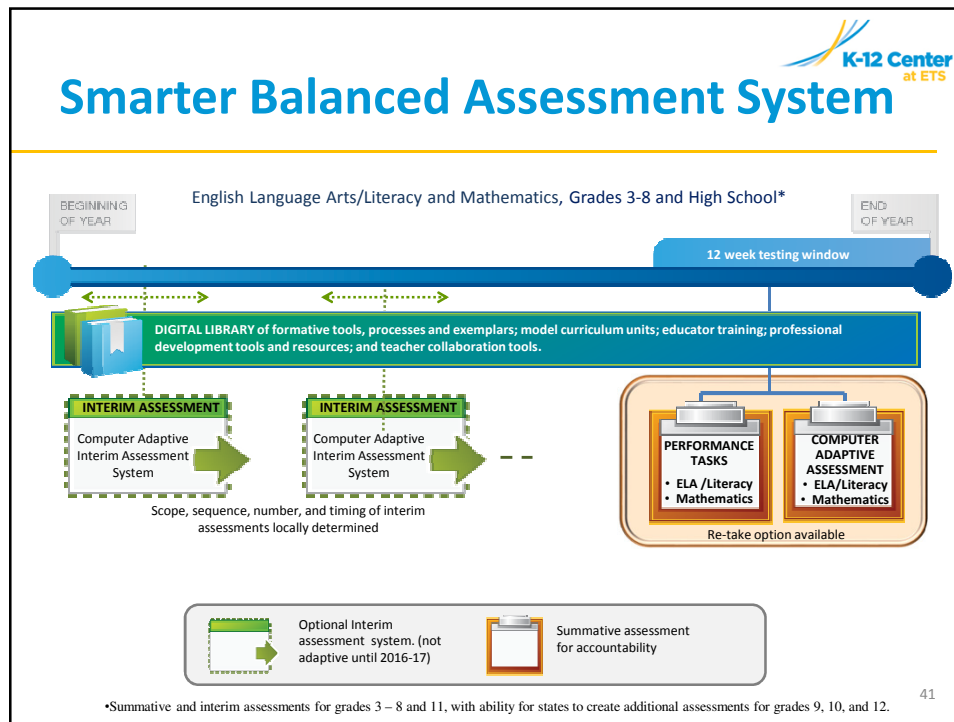
39


Smarter Balanced Supports: The Digital Library



Launched Fall 2014, created “by teachers, for teachers”, with:

- **Vetted teacher materials and commissioned materials**
- **Assessment literacy modules**
- **Formative assessment process descriptions, exemplar instructional modules, vetted and rated instructional resources and professional learning resources**
- **Social media functionality allows teachers to rate materials and share their expertise with educators across the state or across the country**⁴⁰






Smarter Balanced Assessment System

Questions & Discussion

43



Key Similarities

SUMMATIVE ASSESSMENTS

- Online assessments for Grades 3-8 and high school, ELA/literacy and mathematics
- Use of a mix of item types including selected response, constructed response, technology-enhanced and complex performance tasks
- Two components, both given during final weeks of the school year
- Delivery supported on computers, laptops and tablets and a limited variety of operating systems

44



Key Similarities, continued

OTHER ASSESSMENTS, RESOURCES, AND TOOLS

- Online practice tests by grade/course
- Optional diagnostic/interim tests
- Professional development modules
- Formative items/tasks for classroom use
- Online reporting suite
- Digital library for sharing vetted resources and tools
- State ownership and control of all individual student data, as is currently the case for state assessments

45



Key Differences

	PARCC	Smarter Balanced
Grades Tested	3 – 11	3 – 8 and 11 (SEAs can add tests for 9, 10, and 12)
End-Of-Year Assessment	Fixed form	Adaptive
Performance Tasks	3 ELA/Literacy tasks 2 Mathematics testing sessions, multiple tasks	1 ELA/Literacy task 1 Mathematics task
Languages	Translations provided with costs shared by States that need them	Pop-up glossaries and direction translations provided in Spanish, Vietnamese, Arabic, Tagalog/Ilokano, Cantonese, Mandarin, Korean, Punjabi, Russian, Ukrainian and American Sign Language
Retest Policy	State decision, but Consortium will make available 1 retest for grades 3-8 and 3 for HS	State/local decision, but consortium will make available 1 retest option for grades 3-8 and 11 in case of a testing anomaly ⁴⁶



Key Differences, continued

	PARCC	Smarter Balanced
Human Scoring	Vendor trained scorers	State decision on trained educators vs. vendor scoring, with priority placed on teacher involvement, after training
Test Delivery System	Vendor's system (TestNav) thru 2017-18; Then SEAs can manage/contract using PARCC delivery system; open-source system to follow	Free open-source system made available, but states are to either manage or contract for delivery
Use by Non-member States	Non-member states may, for a fee, use PARCC test forms (not items alone), in accordance with security provisions, test administration and reporting requirements, and copyright agreements	Non-member states may gain access to consortium materials, including the item bank, for the same per-pupil fee, in accordance with security provisions, administration and reporting requirements, and copyright agreements

47



Initial Costs

	PARCC	Smarter Balanced
Total Per Pupil Cost for Online Summative Assessments in ELA/Literacy and Mathematics	\$24.00 plus "small project management fee" Includes test delivery and central scoring for 2014-15, Excludes local technology and bandwidth costs	\$22.50 Includes estimated costs for test delivery and scoring, Excludes local technology and bandwidth costs
Paper and Pencil version	Approx. \$33.00	Approx. \$33.50
Additional Per Pupil Cost for Optional System Components	TBD	\$4.80

48



Reporting of Assessment Results: ELA

	PARCC	SMARTER BALANCED
Overall ELA/Literacy Performance Score	5 Levels Level 4 indicates “on track” to college and career readiness	4 Levels Level 3 indicates “on track” to college and career readiness
Reporting Claim Scores	Overall Reading Claim Overall Writing Claim plus 2 sub-claims <ul style="list-style-type: none"> written expression knowledge of language and conventions 	Overall ELA Claim plus 4 sub-claims: <ul style="list-style-type: none"> reading writing listening research/inquiry


49



Reporting of Assessment Results: Math

	PARCC	SMARTER BALANCED
Overall Mathematics Performance Score	5 Levels Level 4 indicates “on track” to college and career readiness	4 Levels Level 3 indicates “on track” to college and career readiness
Reporting Claim Scores	Overall mathematics claim plus 5 sub-claims: <ul style="list-style-type: none"> major content additional and supporting content reasoning modeling fluency (grades 3-6) 	Overall mathematics claim plus 3 sub-claims: <ul style="list-style-type: none"> concepts and procedures problem solving and modeling and data analysis communicating reasoning


50



Next Generation Assessment Systems

Questions & Discussion

51



Presentation: Part 4

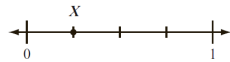
- Key Drivers of Next-Gen Assessments
- The “Call to Action”
- Next Generation Systems of Assessments
- **Comparison of Traditional and Next-Gen State Assessment Items**
- Looking to the Future

52

Next-Gen Assessment Items

Mathematics, Grade 3

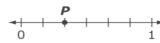
Nina put point X on a number line, as shown below.



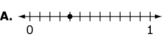
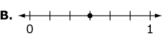
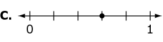

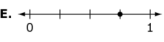
Which fraction best shows where Nina put point X ?

- Ⓐ $\frac{1}{1}$
- Ⓑ $\frac{1}{2}$
- Ⓒ $\frac{1}{4}$
- Ⓓ $\frac{1}{5}$

Look at point P on the number line.



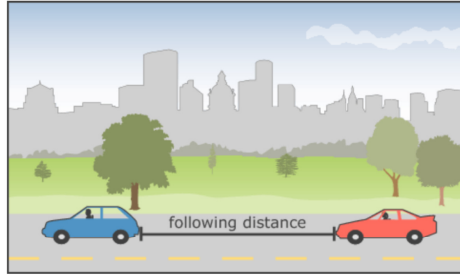
Look at number lines A – E. Is the point on each number line equal to the number shown by P ? Choose Yes or No.

- A.  Yes No
- B.  Yes No
- C.  Yes No
- D.  Yes No
- E.  Yes No

High School Mathematics (Smarter Balanced)



The “two-second rule” is used by a driver who wants to maintain a safe following distance at any speed. A driver must count two seconds from when the car in front of him or her passes a fixed point, such as a tree, until the driver passes the same fixed point. Drivers use this rule to determine the minimum distance to follow a car traveling at the same speed. A diagram representing this distance is shown.



As the speed of the cars increases, the minimum following distance also increases. Explain how the “two-second rule” leads to a greater minimum following distance as the speed of the cars increases. As part of your explanation, include the minimum following distances, in **feet**, for cars traveling at 30 miles per hour and 60 miles per hour.

55

Writing, Grade 6



Typical prompt to date:

Going to the movies is a major source of entertainment for many students. Imagine that the only discount movie theater in your area is closing. Write a persuasive essay in support of keeping the discount movie theater open.

New Task Type: Performance Task (Smarter Balanced)

- Students are taken through a 20 minute Classroom Activity to introduce the topic of school and community gardens. They **read two articles** and **watch a 3 minute 5 second video**. The entire class engages in a short conversation about the articles and video.
- Students are then given 35 minutes to respond to two short constructed response questions. After a break, students are given **70 minutes** to respond to the following prompt:

Some students have suggested that a student gardening program should be started at your school. You are working on the school newsletter, and your assignment is to write an argumentative article ... In your article, you will take a side... Support your position with information from the sources you have examined. The audience for your article will be the teachers and students at your school.

56



ELA Task, High School (traditional)

Typical prompt to date:

Going to the movies is a major source of entertainment for many students. Imagine that the only discount movie theater in your area is closing. Write a persuasive essay in support of keeping the discount movie theater open.

57



ELA Task, High School (PARCC)

Literary Analysis Task

Students read excerpts from two novels, both of which are about the experiences of bi-racial individuals at the turn of the 20th century. Both are “complex texts”, using advanced vocabulary, sentence structure, and literary techniques.

Excerpt from *Quicksand*:

“...Helga stared into the approaching night, glad to be at last alone, free of that great superfluity of human beings, yellow, brown, and black, which, as the torrid summer burnt to its close, had so oppressed her. No, she hadn’t belonged there. Of her attempt to emerge from that inherent aloneness which was part of her very being, only dullness had come, dullness and a great aversion.”

58

ELA Task, High School, cont'd (PARCC)



After several short questions about these excerpts, students are given an extended response task:

"Now that you have read and answered questions about the passages from Quicksand and The Autobiography of an Ex-Colored Man write an essay in which you identify and explain a theme that is similar in both passages. In your essay, discuss how each author uses the characters, events, and settings in the passages to develop the theme."

59

Next Generation Assessment Items



Questions & Discussion

60



Presentation: Part 5

- Key Drivers of Next-Gen Assessments
- The “Call to Action”
- Next Generation Systems of Assessments
- Comparison of traditional and next-gen state assessment items
- **Looking to the Future**

61



Looking to the Future

- Use of advanced assessment development processes (ECD)
- Inclusion of complex items/tasks on state assessments to better measure critical skills
- Utilization of the power of computer-based testing
 - To improve accessibility and validity for students with disabilities and English learners
 - To expand the types and complexity of constructs that can be measured, including more of the critical higher order skills
- States struggle to balance need to assess complex competencies with reasonable testing time and cost⁶²



www.k12center.org

Thank you.

Pascal (Pat) D. Forgione, Jr., Ph.D.

Distinguished Presidential Scholar and Executive Director
Center for K-12 Assessment & Performance Management at ETS

E-Mail: pdforgione@k12center.org

Driving Advances in K-12 Assessment