Standards-Based Competency Assessments, Grades 3-11

Arizona Department of Education

TECHNICAL ORIGINAL

RFP#: ADED14-0003
DUE: December 5, 2013

ADE
c/o Chief Procurement Officer,
Lupita Gomez
Procurementinbox@azed.gov

ALWAYS LEARNING

PEARSON

PEARSON ALWAYS LEARNING

ASSESSMENT & INSTRUCTION
State Services
NCS Pearson, Inc.
2510 N Dodge St
Iowa City, IA 52245

T: 319.339.6620
F: 319.339.6477
M: 319-621.4042
E: Rich.Young@Pearson.com

www.pearson.com

Lupita Gomez
Chief Procurement Officer
Arizona Department of Education
1535 W. Jefferson St.
Phoenix, AZ 85007

RE: Request for Information ADED14-0003 Standards Based Competency Assessment, Grades 3-11

5 December 2013

Dear Ms. Gomez:

Starting in the 2014–15 school year, the Arizona Department of Education (ADE) will begin to replace Arizona's Instrument to Measure Standards (AIMS) with next-generation assessments that will provide parents and educators throughout the state with rigorous new measures of student achievement, a higher bar for achieving excellence, and a more unified focus on preparing all students for college and careers. Through these actions, Arizona will take another significant step toward completing the cycle of reform initiated by Governor Brewer and her P-20 Coordinating Council (now the Arizona Ready Council) more than four years ago.

Pearson's history of support for Arizona's assessment programs stretches back to the very beginning of the AIMS program more than 15 years ago, and it has been an honor and a privilege to have played such an important role in Arizona's growth and success since that time. In fact, with more than 650 full-time and 1,500 part-time Arizona-based employees, many of whom are products of (and whose children attend) Arizona schools and post-secondary institutions, we view it as our *responsibility* to do all we can to help Arizona succeed.

Throughout our 90-plus years in educational assessment, we have helped many other states through transitions similar in magnitude and impact as the one Arizona has now undertaken, providing consistent leadership and innovation through technology and advanced measurement research. Achieving the gains Arizona now seeks will require new assessments that reflect the increased rigor of Arizona's standards, leverage technology to provide deeper insight into students' knowledge, skills, and abilities, and ultimately serve to illuminate each child's path to college and career readiness.

As you prepare your Request for Proposals (RFP) to select Arizona's next assessment system, the ADE will have a wide range of options from which to choose and, as a result, a number of very important decisions to make. While the use of technology in next generation assessments allows for richer, more authentic assessment, computer-based testing presents an entirely new set of practical and logistical challenges for schools, especially in states like Arizona where high-stakes online testing is new. Testing systems must work with schools' existing hardware and software environments, be easy for local technology support personnel to set up and manage, and be easy for students and teachers to use. In addition, schools or districts that lack sufficient hardware, network, or staff capacity for full online testing implementation will need the flexibility to continue testing on paper as their capabilities and experience with computer-based testing grow over time.

Having implemented high-stakes online testing in thousands of schools and districts in more than a dozen states over the past 13 years, we have the knowledge, experience, and insight it takes to successfully navigate these challenges. Our PearsonAccess and TestNav platforms have been specifically designed to meet the unique requirements of both paper and online testing, offering schools and districts the flexibility to simultaneously manage the demands and meet the requirements of testing in both modes. Options we can deliver or support for Arizona include:

- **The Partnership for Assessment of Readiness for College and Career (PARCC).** Schools and districts throughout Arizona will be administering PARCC field tests through PearsonAccess and TestNav in spring 2014. These same systems will also be used for the first operational administration in spring 2015.

- **ACT Aspire™.** A new assessment system developed jointly by ACT, Inc. and Pearson to measure growth to college and career readiness from grade 3 through 10 in reading, writing, mathematics, and science. With the optional ACT Assessment serving as its capstone at grade 11, ACT Aspire can also be combined with ACT's Quality Core assessments to provide additional alternatives for measuring readiness through end-of-course achievement.

- **Custom Assessments.** Using portions of Arizona's existing item bank in combination with next-generation items licensed from Pearson or one of the two national assessment consortia, Arizona could create custom assessment blueprints to incorporate a mix of item types that would be more efficient and less costly to create, administer, and score. These assessments could then be "ratcheted up" as experience, capabilities, and available resources increase over time.

To remain competitive, Arizona students must keep pace with their peers across the country and around the world. To facilitate their growth and progress, Arizona's assessment systems must keep pace as well. Pearson can provide Arizona with the stability it needs to move forward with confidence, and the advanced research and technology it takes to accelerate growth and achievement for all Arizona students.

Regardless of the path Arizona chooses, Pearson will be there to help lead the way. I have actively followed the evolution of Arizona's education system for nearly a decade now and have had the privilege of personally supporting the success of your current assessments over the last five years. As Arizona continues to advance its education reforms, I look forward to the opportunity to continue serving the ADE any way I can. If you have any questions about the information we have provided in our enclosed response or need further information of any sort, please do not hesitate to contact me at 319.621.4042 or via email at Rich.Young@pearson.com, and I will be glad to assist you as quickly and efficiently as possible. Thank you for your consideration.

Sincerely,

Rich Young
Vice President, State Services
Assessment & Instruction
Pearson

REQUEST FOR INFORMATION

RFI # **ADED14-0003**

Description: **Standards-Based Competency Assessments, Grades 3 - 11**

Due date: **December 5, 2013** **Time: 3:00 p.m. (mst)**

General Information

In accordance with ARS §41-2555, the State Board of Education (SBE) through the Arizona Department of Education (ADE) hereby gives notice via this Request for Information (RFI), of our desire to receive responses, feedback, suggestions and comments regarding the items set forth herein.

In accordance with R2-7-G301, responses to the RFI are not offers and cannot be accepted to form a binding contract. To the extent allowed by law, information contained in a response to an RFI may be considered confidential until the procurement process is concluded or two (2) years from when the RFI's due date has elapsed, whichever occurs first. Responders are solely responsible for expenses associated with responding to the RFI. The State is under no obligation to the responders in so far as the next steps to this process are concerned.

Responses to the RFI shall be submitted as follows:

One (1) e-mail copy of the informational documents, in PDF format, shall be submitted to ADE, c/o Chief Procurement Officer, Lupita Gomez; Procurementinbox@azed.gov by December 5, 2013 at 3:00pm, MST.

SBE is issuing this Request for Information (RFI) to solicit data from vendors of educational assessments that could be used in school year 2014-15 as standards-based competency assessments for grades 3 – 11.

Background:

A.R.S. § 15-741 requires the SBE to adopt and implement a criterion-referenced assessment to measure pupil achievement in the board adopted academic standards in reading, writing and mathematics. In addition to meeting federal testing requirements under the Elementary and Secondary Education Act, this assessment is used for state mandated accountability measures, including Arizona's school accountability system (A.R.S. 15-241), teacher and principal evaluations (A.R.S. 15-203), and third grade retention (A.R.S. 15-701).

In June 2010, the SBE adopted new English Language Arts and Mathematics standards designed to demonstrate college and career readiness and aligned with the Common Core Standards. The implementation of these new standards was phased in over several years, and the new standards will be tested for the first time in school year 2014-15.

The SBE issues this RFI in order to better understand the assessment options that may be used as a statewide standards-based competency assessment for grades 3-11 starting in 2014-15. SBE welcomes responses for both existing assessments, and those under development that will be ready for implementation in 2014-15. The information gathered will help inform the SBE as it further develops a Request for Proposal (RFP) to procure a final assessment for use in Arizona.

Information being requested:

This RFI requests that vendors provide brief descriptions and cost estimates for any existing vendor products or systems or any customized products they have developed with state(s) or other entities which meet the criteria of one or more of the components described below. This RFI is being issued to obtain information only and is not intended to result in contracts or vendor agreements with any respondent. Respondents to this RFI will not be excluded from eligibility to participate in any future requests for proposals (RFPs) based on their decision to respond to this RFI or on the content of their responses to this RFI.

While we are interested in responses that encompass all functional components, we are also interested in responses that cover discrete functional components of the entire system.

Please provide answers to the questions in the prescribed format and supporting materials that describe the proposed assessment and related issues.

1. **Background**
   a. Provide a brief history of the organization and its governance structure.
   b. Identify the individuals from the organization that will be working with Arizona officials on all aspects of the assessments' implementation.

2. **Overview of assessment**
   a. Describe the assessments, the grades assessed, the subject areas included, and the formative, diagnostic and summative components. Also describe available end-of-course assessments for grades nine through eleven.
   b. Describe the timeline for the development of the assessments to ensure full implementation by the 2014 – 2015 academic year.
   c. Describe how, and the degree to which, the assessments are specifically aligned to Arizona's academic standards in mathematics and English language arts (reading and writing), include any alignment studies, if available.
   d. Describe how the assessments' results can be compared to other states' criterion-referenced assessments expected to be in use beginning in the 2014-2015 academic year.
   e. Describe how the assessments are aligned to college / career expectations. Describe the validation process, including the role of post-secondary education in establishing the readiness expectations. Include any alignment studies, if available.
   f. Describe the available accessibility features, as well as assessment accommodations for individuals with disabilities and English Language Learners.
   g. Describe any practice and/or sample assessments that are available.

3. **Computer-based assessments**
   a. If applicable, describe the computer-based option(s) for the administration of the assessments.
   b. If applicable, describe the technological specifications for the administration of the assessments. This should include specifications for computer hardware, input devices, security requirements, bandwidth, web browser requirements, and platform software.

c. Describe any available training and technical support that is available for the computer-based assessments.

d. If the assessment is computer-based is there a paper / pencil option?

4. Assessment administration

a. Describe the total anticipated testing time for each assessment (mathematics, reading, and writing) by grade level. If computer-based, include the calculated student to device ratio.

b. Identify the anticipated testing window for each assessment by grade level.

c. Describe the training needs and available training for teachers and administrators pertaining to the administration of the assessments.

5. Assessment standards setting and scoring

a. Describe the standards setting process.

b. If already established, describe the performance levels and the performance level descriptors for each category.

c. Describe the score reports available to teachers, students and parents. The description should include:

   i. How the reports illustrate a student's progress on the continuum toward college and career readiness, grade by grade, and course by course; and

   ii. How the reports are instructionally valuable, easy to understand by all audiences, and are delivered in time to provide useful, actionable data to students, parents, and teachers.

d. Describe the process and timelines for scoring the assessments. Include computer-based and pencil / paper processes and timelines, as applicable.

e. Describe how scores on the assessments will be comparable to other common college/career ready assessments.

6. Assessment development

a. Describe how the development of the assessments will adhere to the principles of universal design, so that the testing interface, whether paper- or technology-based, does not impede student performance.

b. Describe any comparability studies between the paper/pencil and computer-based assessments.

c. Describe the processes for item development. The description should include:

   i. How the reading and writing items will require students to demonstrate a range of higher-order, analytical thinking and performance skills in reading, writing and research based on the depth and complexity of the standards, allowing robust information to be gathered for students with varied levels of achievement; and

   ii. How the mathematics items will require students to demonstrate a range of performance based on the depth and complexity of the standards, allowing robust information to be gathered for students with varied levels of achievement.

d. Describe the procedures used to ensure all test items are properly aligned to applicable standards and avoid bias. Include the role of state representatives in these processes.

e. Describe the procedures for ongoing assessment item development, including procedures to accommodate any future modifications to the relevant academic standards.

f. Describe field testing procedures and timelines for the assessments. Include information regarding field testing of Arizona students, if applicable.

7. Assessment Cost

a. Describe the pricing structure including the cost per student.

Instructions

Any submitted materials considered by the respondent to be proprietary or confidential shall be clearly marked as such and provided as a separate section in your response. In response to such requests, the State will make reasonable efforts to maintain confidentiality of these materials to the extent permissible by law. However, should the State be required to release these materials, the State will provide the respondent reasonable notice in advance of doing so in order to allow the respondent time to take any action to prevent these materials from being released.

CLARIFICATION OF SOLICITATION REQUIREMENTS

It is the responsibility of all Offerors to examine the RFI and seek clarification of any item or requirement that may be clear or unclear to them and to check all offers for accuracy before submittal to SBE/ADE.

Any questions regarding this solicitation can be answered by emailing the contact person listed on the title page of this solicitation. All questions must be in writing. The Chief Procurement Officer is the only authorized person to give information as to the requirements of the solicitation in addition to that contained in the written documents.

Questions should be emailed and received by Chief Procurement Officer no less than seventy-two hours before the RFI due date and time.

All questions must reference the page and item that is to be addressed.
All addendums should be acknowledged using the Addendums Acknowledgement Form
included in this document. No verbal or written information, which is obtained other than by information in this document or by Addendum to this solicitation, will be binding on the Arizona Department of Education.

I/We Rich Young the undersigned, do hereby submit this response for information with regard to the Standards-Based Competency Assessments in accordance with ARS §41-2555 Request for Information specification contained herein.

	12/4/2013
Name	Date
NCS Pearson, Inc.	VP, State Services
Company Name	Title
2510 North Dodge St. Iowa City , IA 52245	319.339.6620
Address	Telephone Number

Table of Contents

Executive Summary **ES–1**

Background **1–1**

Overview of Assessment **2–1**

Computer-Based Assessments **3–1**

Assessment Administration **4–1**

Assessment Standards Setting and Scoring **5–1**

Assessment Development **6–1**

Assessment Cost **7–1**

Executive Summary

Since the origins of Arizona's first Race to the Top application in 2009, Governor Brewer's education reform agenda has consistently focused on college and career readiness as the ultimate standard by which success for all Arizona students should be defined and measured. By establishing rigorous, high-quality academic standards and assessments as one of Arizona's four pillars of reform, Governor Brewer has established high expectations for achievement that educators throughout the state are working hard to fulfill.

Unified in their support for Governor Brewer's reform agenda, Superintendent Huppenthal, the Arizona Department of Education (ADE), and the Arizona State Board of Education have undertaken a series of bold actions to implement some of the most fundamental changes to Arizona's education landscape since Arizona's Instrument to Measure Standards (AIMS) was introduced more than 15 years ago. Now, with Request for Information (RFI) ADED14-0003 for Standards-Based Competency Assessments Grades 3–11, the ADE has entered the final stages of Arizona's transition to a new assessment system, a process that has been more than four years in the making.

To inform the ADE's development of an RFP that fulfills Arizona's original reform vision, this RFI seeks important background information on reading, writing, and mathematics assessment options available for implementation in 2014–15 that offer the following high-level features:

- Alignment with Arizona's College and Career Ready Standards
- Options to create a more comprehensive and balanced system of formative, interim, diagnostic, end-of-course, and summative assessments
- Options for both paper-pencil and computer-based administration
- Comparisons to college and career readiness indicators and measures in other states
- Accessibility and accommodations for students with disabilities and English language learners
- Reporting options to help parents and educators better understand their students' needs and progress toward readiness

To better prepare students for success in an increasingly competitive, global economy, states across the nation are implementing innovative new next-generation assessment systems that encompass these key features and more. However, with so many changes and so much riding on Arizona's assessment system, failure, as they say, is *not* an option, and parents, teachers, administrators, and support staff are understandably anxious about the many unknowns that lie ahead.

Fortunately, as the ADE's current contractor for both the AIMS and Arizona's English Language Learner Assessment (AZELLA), Pearson can help deliver or support a variety of future assessment options that

fulfill Arizona's vision for a next-generation, college and career readiness assessment system while providing your stakeholders with a greater sense of comfort, confidence, and stability. These options include:

- **The Partnership for Assessment of Readiness for College and Career (PARCC).** Schools and districts throughout Arizona will be administering PARCC field tests in spring 2014 using TestNav™, our proven online delivery platform, and PearsonAccess™, the same online administrative portal used to manage AIMS and AZELLA. These same systems will also be used for the first operational PARCC administration in spring 2015, providing Arizona with a high degree of familiarity and continuity.
- **Assessments from ACT, Inc. and ACT Aspire™.** ACT Aspire is a new assessment system developed jointly by ACT, Inc. and Pearson to measure growth to college and career readiness from grade 3 through 10 in English, reading, writing, mathematics, and science. With the optional ACT assessment serving as its capstone at grade 11, ACT Aspire can also be combined with ACT's QualityCore® assessments to provide additional alternatives for measuring readiness through end-of-course achievement.
- **Custom Assessments.** Using portions of Arizona's existing item bank in combination with next-generation items licensed from Pearson or either of the two national assessment consortia, Pearson can help Arizona create custom assessment blueprints that incorporate a mix of item types designed to be more efficient and less costly to create, administer, and score. If desired, these assessments could be further "ratcheted up" to provide increased rigor and depth as experience, capabilities, and available resources increase over time.

With each of these options, Pearson offers a common thread that Arizona can rely on, with our proven online technology, large operational capacity, comprehensive program management, and track record of success in supporting Arizona through other high-stakes transitions in the past. Coupled with our experience developing and delivering next-generation college and career readiness assessments in other states like Texas, New York, Kentucky, and Virginia, the ADE can count on receiving superior service and support from Pearson regardless of the role we play.

Our Support for PARCC

As a governing state in the PARCC consortium, Arizona is well familiar with the features and benefits of the PARCC assessments. Pearson is strongly committed to PARCC's success and is currently supporting the consortium through prime contracts for the Technology Readiness Tool, item development, and online test delivery and scoring. Arizona schools and districts participating in the spring 2014 paper and online field tests will have an opportunity to experience firsthand the PARCC assessments and Pearson's delivery support for PARCC, with many schools participating in online testing for the first time.

If Arizona should choose to continue with the PARCC assessments in 2014–15, Pearson will be on board to support a smooth transition and provide continuity to Arizona districts and schools for online test delivery. In addition, we continue to pursue opportunities that would expand our service to PARCC beyond online delivery in spring 2015. However, due to the ADE's potential role in PARCC procurements, it would be inappropriate for us to further elaborate on the PARCC assessments or speculate on the potential scope of Pearson's future support for PARCC in this response.

Teaming with ACT, Inc. and ACT Aspire

The digital revolution is well underway in Arizona classrooms, so to achieve the kind of gains Arizona now seeks, the ADE needs an assessment system designed to support and promote the changes in instruction now taking place. With college and career readiness as Arizona's ultimate goal, no other company is as qualified to provide Arizona's next assessment system as ACT, whose flagship ACT assessment has served as the gold standard for college entrance nationwide for more than 50 years. Throughout its history, ACT's research-based programs and services for educational assessment, workforce development, career planning, and professional certification and licensure have become mainstays of the American educational and business landscapes.

Like Pearson, ACT is deeply involved in supporting Arizona's education system, through innovative programs like the Arizona ACT Initiative sponsored by the Governor's Office of Education Innovation and Helios Education Foundation, and the Gaining Early Awareness and Readiness for Undergraduate Program (GEAR UP), a college access program offered through Northern Arizona University designed to significantly increase the number of students from low-income communities who stay in school and are prepared to succeed in postsecondary education.[1]

With common origins at the University of Iowa dating back to the 1950's, [redacted] Through our most recent collaboration, ACT and Pearson have formed a new entity, ACT Aspire, LLC, to create ACT Aspire, a next-generation assessment designed to measure growth toward college and career readiness for students in grades 3–10, with Work-Ready and ACT Composite prediction scores for students in grades 9 and 10.

Launching in spring 2014, the ACT Aspire assessment system includes summative, diagnostic, and classroom-based assessments that can be used in combination with The ACT as its capstone for grade 11 and with ACT QualityCore assessments for measuring end-of-course achievement. [redacted]

Options to Create a Custom Assessment

If Arizona should determine that a shared or pre-built assessment will not fully meet the state's needs, the ADE could procure a custom assessment built to a blueprint of Arizona's own design, using a combination of high-quality, well-aligned AIMS items and items licensed from PARCC, the SMARTER Balanced Assessment Consortium, and/or Pearson. [redacted]

[1] Citation: https://nau.edu/COE/About/Projects/GEAR-UP/. © 2013 Arizona Board of Regents. Northern Arizona University, Flagstaff, Arizona

Depending on the source(s) of items selected, the ADE and Pearson, working collaboratively together as we have in the past, would likely be able to assemble and deliver custom operational paper and online forms by spring 2015. In addition, there would be a strong likelihood that custom assessments built from these sources would be able to provide equivalent measures for comparability to other states' results. Again, with solid operational support and proven technology from Pearson, this approach would provide for reduced risk compared other custom options and a smooth transition for all Arizona stakeholders.

Proven Delivery, Powered by Pearson

In addition to the significant curricular and instructional changes taking place throughout the state, one of the biggest challenges Arizona faces is the transition to online testing. Over the past decade, Pearson has implemented more than 20 online assessment programs for state and national customers, including current customers in Colorado, Florida, Georgia, Maryland, Mississippi, Tennessee, Texas, and Virginia. Our proven technology and ability to successfully manage states' transitions from paper to online testing, coupled with our experience and established relationships throughout Arizona, make us well qualified to help lead your program into the future of next generation assessment.

When evaluating an online test delivery system, Arizona should consider five basic criteria to determine which system is the right fit for its assessment program:

- **Security** to protect test content and prevent cheating
- **Consistency** for students to receive a fair and comparable testing experience, regardless of the hardware, software, or device they use
- **Scalability** to complete statewide testing efficiently and on time
- **Reliability** to protect students from disruptions and the potential for lost work, allowing them to focus on performing to the best of their abilities
- **Total Cost of Ownership** to reduce the local burden of preparing for and supporting testing through zero footprint, industry-standard software solutions

With tens of millions of high-stakes online assessments delivered over the last decade, more than 11 million so far this year, Pearson's TestNav and PearsonAccess online platforms offer Arizona a proven solution that fulfills all of the above requirements. For each of the assessment options detailed throughout our response, Pearson provides online delivery, operational capacity, and program management support in subcontractor, partnership, and prime contractor roles, affording ADE the comfort of knowing that no matter which route you choose, your trusted ally can continue to provide Arizona with strong support and dependable service.

Additional Considerations

As you prepare your RFP, we would like to offer some additional considerations based on our experience in other states around the country. One such consideration is to create or maintain a more cohesive overall assessment system that encompasses all subject areas and even other types of assessments. For example, other states undergoing similar transitions are requiring that new assessments for core subjects in English language arts (ELA) and mathematics continue to be closely coordinated with legacy programs

as well as new or existing assessments in science or other subject areas. By taking a more holistic, comprehensive approach to reform, the ADE can help educators make significant advancements in foundational skills while continuing to support the importance of other subject areas. This type of integrated approach can help promote greater cross-curricular literacy, a key tenet of Arizona's new ELA standards. Maintaining a more cohesive and unified assessment system also helps reduce the overall administrative burden on schools, making it less disruptive to introduce other incremental reforms over time.

Another potential consideration pertains to requirements for integration with state data systems. Given the ADE's ongoing initiative to implement the Arizona Education Learning and Accountability System (AELAS) and Arizona Education Data-driven Decision System (AzED^{3}S), your RFP for a new assessment system could include specifications for operational synergies that would likely be more costly and time-consuming to add later.

Thank You

With deep roots in Arizona's public education system and more than 2,000 full- and part-time Arizona-based employees, Pearson is deeply invested in your success. Regardless of what the future may bring, as citizens of Arizona's business and education community, we view it as our responsibility to do everything within our power to see that all Arizona children have a chance to succeed on their path to college and career readiness. We pledge our support to this important undertaking, and we look forward to continuing our successful relationship with you. Thank you for providing us with this opportunity.

1. Background

Requirement

a. Provide a brief history of the organization and its governance structure.

Response

About Pearson

As the leading education services company, Pearson is serious about evolving how the world learns. We apply our deep education experience and research, invest in innovative technologies, and promote collaboration throughout the education ecosystem. Real change is our commitment and its results are delivered by connecting capabilities to create actionable, scalable solutions that improve access, affordability, and achievement.

Pearson currently provides large-scale assessment services in more than half of US states as well as for the US Department of Education, the College Board, the National Board for Professional Teaching Standards, and for the Partnership for the Assessment of College and Career Readiness (PARCC).

Pearson contracted with the Arizona Department of Education (ADE) in 1998 to create Arizona's Instrument to Measure Standards (AIMS) and helped manage the program's evolution through 2004, in compliance with the No Child Left Behind Act of 2001. In 2007, Pearson conducted a special study for the AIMS Grade 8 science assessments, comparing results from four field test forms administered both on paper and online through Pearson's TestNav system. Results showed no statistical difference in student results for tests administered in either mode.

In 2013, Pearson collaborated with the ADE to introduce college and career readiness target scores into AIMS score reports. These targets, established for Arizona by the National Center for Educational Achievement (NCEA, a department of ACT, Inc.), serve as grade-level indicators of whether students are on target to graduate from high school ready for college and career. Pearson has continued to serve as the ADE's AIMS contractor since 2009.

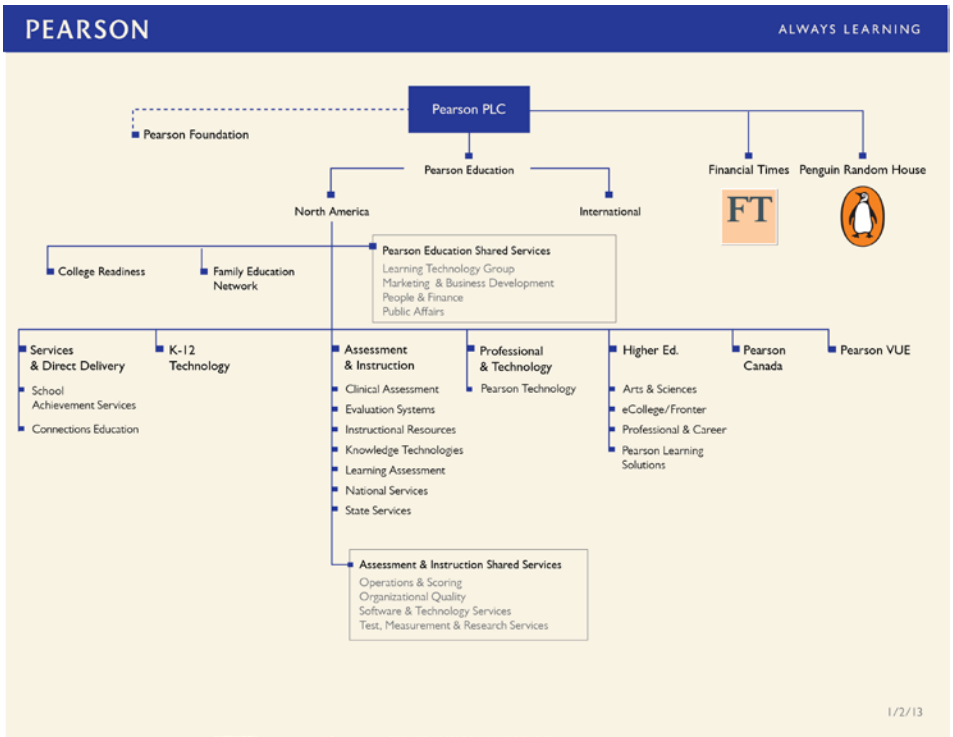
In 2004, the ADE contracted with Pearson to implement Arizona's first statewide English language proficiency assessment, the Stanford English Language Proficiency which was subsequently customized to better align with Arizona's English Language Proficiency (ELP) Standards and renamed the Arizona English Language Learner Assessment (AZELLA). In 2011, Pearson teamed with WestEd to create a third edition of the AZELLA based on Arizona's new ELP Standards, which were revised in partnership with WestEd to align with the academic expectations of Arizona's College and Career Ready Standards.

With the new AZELLA, Arizona has successfully satisfied stringent new federal requirements and become the first state in the country to implement standardized, technology-based speaking assessments using artificial intelligence (AI) scoring.

A World of Resources from Pearson

Pearson's three main lines of business represent three key stages of learning: school, higher education, and professional. These lines of business serve markets around the world with a wide range of high-quality products and services in assessment, instruction, technology platforms, and educator services.

Based in London, Pearson employs approximately 36,000 people in more than 60 nations around the world, including more than 650 full-time and 1,500 part-time employees in Arizona.

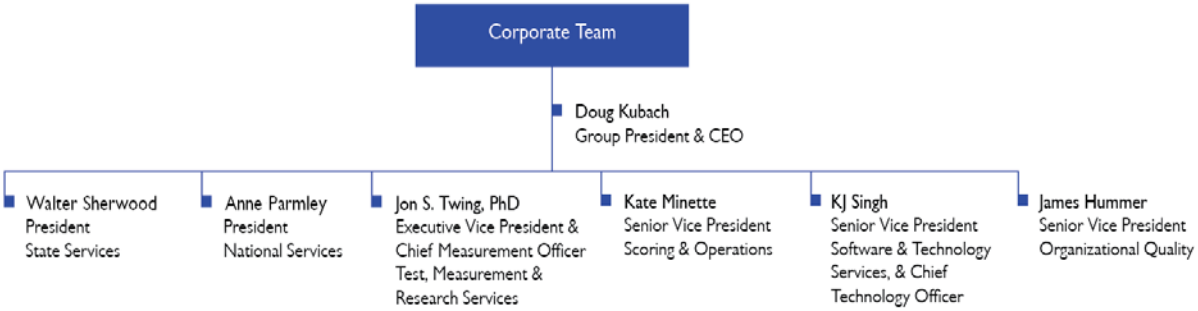


Pearson Organizational Chart. The State and National Services groups under Assessment & Information are part of a suite of education companies comprising Pearson Education.

Douglas Kubach is President and CEO of Pearson's US Assessment and Instruction (A&I) group, a business unit of NCS Pearson Inc., of Bloomington, MN (incorporated in 1962).

NCS Pearson Inc.
5601 Green Valley Drive
Bloomington, Minnesota 55437

NCS Pearson Inc. is owned by PN Holdings Inc., which is owned by its ultimate parent, Pearson plc (a publicly held U.K. company), or its subsidiary(ies).



Senior Leadership. Pearson's senior corporate leadership team has experience and diverse skills to oversee the success of Arizona's assessment programs.

The A&I leadership team represents each key area of our large-scale assessment business, setting high standards for quality, timeliness, and integrity in the measurement, operations, and technology services Pearson delivers for Arizona.

Insights for Innovation

Expectations for educators and students have never been higher. Students today must be prepared for a dynamic, global economy—one that calls on people to use data more thoughtfully, think creatively, and work productively in teams. Pearson's Research & Innovation Network (researchnetwork.pearson.com) works toward achieving these goals in four ways:

- **Inventing** new digital technologies to support educators in adapting instruction to individual student needs
- **Translating** research into recommendations and services that help educators overcome instructional challenges
- **Collaborating** with leading experts and practitioners to conduct research and development
- **Sharing and receiving** intellectual insights during conferences, through publications, when collaborating with educational stakeholders, and by consulting with foundations on the innovations needed as education navigates its way from the pre-digital to the digital

The Research & Innovation Network comprises five centers with distinct, but related, research agendas including the following:

- **The Center for NextGen Learning and Assessment** conducts research on innovative new item types, performance assessments, and scoring methods to help understand students' day-to-day progress and accelerate their learning.
- **The Center for Online Learning** studies teaching and learning through virtual, blended, adaptive, and other eLearning technologies through multiple devices to identify effective practices and recommend strategies for personalized, digital instruction at scale.
- **The Center for Digital Data, Analytics, and Adaptive Learning** explores new ways to use the massive amounts of data generated by online learning, gaming, and other educational technologies to assess, enable, and personalize learning without the disruption of traditional tests.
- **The Center for College and Career Success** helps develop new ways to define, measure, and track students' readiness to enter college or the globally competitive workforce.
- **The Center for Educator Effectiveness** investigates and promotes new methods to measure teachers' performance, develop their full potential, and achieve greater educational outcomes for their students.

Each center is led by accomplished researchers who both provide intellectual leadership in the field and advise practitioners working at the front lines of education. Their work constantly informs and enhances the services we offer and provide our customers with additional opportunities for collaboration to advance their education reform programs.

A History of Advancing Educational Assessment

As the nation's most comprehensive provider of educational assessment products, services, and solutions, Pearson has been a trusted collaborator in district, state, and national assessments for 90 years. We help students, adults, educators, and professionals use assessment, research, and innovative technologies to promote learning and personal development while advancing academic achievement.

> *Pearson has been a trusted collaborator in district, state, and national assessments for 90 years.*

New technologies have increased collaboration within and among states just as economic pressures to remain globally competitive drive efforts to make learning more effective, accessible, and personalized. In this dynamic environment, the challenge is to design comprehensive, new assessment systems that accelerate educators' ability to personalize learning for every student.

Leading the Way in Next Generation Assessment

Innovative approaches to item development and scoring, such as the expanded use of performance tasks, computer adaptive testing, and automated scoring, are allowing us to more effectively measure what students know and can do. These next-generation assessments move beyond snapshots of student achievement to capture complex performance and track student growth over time, leveraging new digital technologies to provide richer, more timely feedback to teachers, students, parents, and policymakers.

Expert Test Design, Development, and Measurement Services

At Pearson, test development takes place within our Test, Measurement, and Research Services (TMRS) group rather than across several departments. Having our content experts and psychometricians in the same department facilitates communication, project planning, and close collaboration for the most efficient approach to test development.

Our established processes allow our staff to develop items or work with educators to develop items that align to applicable content standards, adhere to the principles of Universal Design, and reflect best instructional practices.

At Pearson, our content development and psychometric research capabilities come together to provide the following:

Pearson Content Development and Psychometric Research Capabilities	
Multiple-choice, constructed response, and technology-based innovative item development	Question and Test Interoperability (QTI) and Accessible Portable Item Protocol (APIP) standards
Test blueprints	Pilot, field test, and operational test design, sampling, and data analysis
Statistical analyses	Validity and reliability investigations
Item development aligned to state standards	Differential Item Analysis (DIF) and other statistical data analyses, such as item and scaling drift
Item reviews for content, bias, and universal design	Standard setting
Test form development, including forms adapted for online administration	Accommodated testing of English as a second language populations and examinees with disabilities
Student/device interaction, usability and comparability studies	Performance and automated scoring studies
Automated item banking, tracking, and test forms construction	Preparation of technical documentation and reports
Calibration, scaling, and equating using a broad array of classical and IRT-based measurement models	Presentations of technical data in a variety of media and to various audiences, including state boards of education and state legislators
Special studies to examine various assessment issues (e.g., comparability studies, generalizability analyses, factor analytic studies, cognitive labs)	Support for states as they defend and explain how they are meeting requirements of Federal Peer Review and measurement best practices

Developing State Assessments. Pearson psychometricians and content development specialists support our state customers through use of sound methodologies in test development.

Dr. Jon S. Twing, Pearson's Chief Measurement Officer, directs three functional teams:

- **Measurement Services**, led by Dr. Walter "Denny" Way
- **Research and Innovation Network**, led by Kimberly O'Malley
- **Learning Integration**, led by Dean Brown

Measurement Services

Arizona's College and Career Ready Standards provide guidance about what students should understand and be able to do at critical points in their education.Developing tests that fairly and reliably measure student learning for these new standards takes experience and capacity—hallmarks of Pearson capabilities. Our item development work for the PARCC consortium has added greatly to our experience, insight, and ability to develop high-quality, innovative new items aligned with college and career ready expectations, offering Arizona a unique advantage.

To support item and test development and test administration in a changing educational landscape, Pearson provides customers with the content and psychometric support shown in the following figure.

Measurement Services for Pearson State Customers	
Initial evaluation of test items and prompts, including alignments	Accommodated versions of an assessment
Content development in all major subject areas, including general population, English language learners, and special needs students	Quality control activities to maintain consistency across item/test form development, administration, scoring, and reporting functions
Incorporation of stimuli, items, and prompts into pilot and field test forms	Development, maintenance, and support for test security measures
Participation in internal and external item/prompt content, bias, and data review	Securing of permissions, payment on behalf of clients, tracking, and records handoff
Item/prompt scoring and subsequent analyses of pilot test data	Descriptions of test administration procedures
Test form development, including descriptions of test blueprints and alignment to clients' learning standards, practice forms, and descriptions of their contents	Training and other consultation

Supporting Assessment Quality. Our established processes enable our staff to develop items that align to client content standards, adhere to the principles of Universal Design, and reflect best instructional practices.

To give each student the best possible opportunity to demonstrate his or her full range of knowledge and skills requires a full range of Pearson personnel—including artists, content specialists, researchers, copyeditors, and fact checkers. Many of our staff hold teaching certifications and have experience teaching in K–12 classrooms. Their commitment to quality education shows in the care they give to creation, review, and production of every test item.

This care results also from established processes developed over decades of assessment experience. Our procedures enable Pearson content specialists to develop items that align to client content standards, adhere to the principles of Universal Design, and reflect best practices, including the 1999 AERA, APA, and NCME Standards for Educational and Psychological Testing. Applicable APA Standards guide test development.

Psychometric Services

Sound methodologies and continuing research lead to assessment results that students, parents, educators, and stakeholders can trust. Pearson psychometricians, statisticians, and research personnel use sound methodologies in technical areas of test development. We are careful to consider the content, format, and context in which tests are given, as well as the potential consequences of using the test to make judgments about students and instructional practices.

Many of our psychometricians are nationally recognized experts in high-stakes standards-based assessment. They participate in review and publication of peer research, and they have been consulted on revisions to the APA standards. Dr. Walter "Denny" Way, Senior Vice President of Psychometric and Research Services, served as a member of the joint committee that revised the AERA, APA, and NCME Standards for Educational and Psychological Testing. The APA Standards form the backbone for sound psychometric practice in educational and psychological testing, promote the sound and ethical use of tests, and provide a basis for evaluating the quality of testing practices.

Psychometric Services for Pearson State Customers	
Descriptions of and rationales for the use of Item Response Theory (IRT)-based models for equating and scaling	Descriptions and results of standards validation activities, including discussion and evaluation of cut scores and results and recommendations of that review, as well as data and results related to continued statistical monitoring of effectiveness of cut scores based on operational data
Equating and scaling of test forms, including any horizontal and vertical equating (i.e., year-to-year and vertical scaling) methods used and equating to control for score drift	Descriptions and results of standard setting activities, including process, sample materials, procedures followed, agenda, schedule, list of participants, data and results, and recommendations for new cut scores resulting from the process
Analysis of item- and test-level data, including classical and IRT-based statistics, DIF analyses, estimates of model fit, item and test characteristic curves, and test information functions	Design, conduct, and document results of validity and reliability studies including, where applicable, system validation and other studies, inter-rater reliability investigations, and standard error estimates associated with student scores
Statistical analyses and expert evaluations of bias for field tested items	Sampling plans for pilot testing, early return of operational test data, and special studies
Standard setting using various models appropriate for the given assessment(s)	Estimates of incidence of misclassification across multiple performance levels

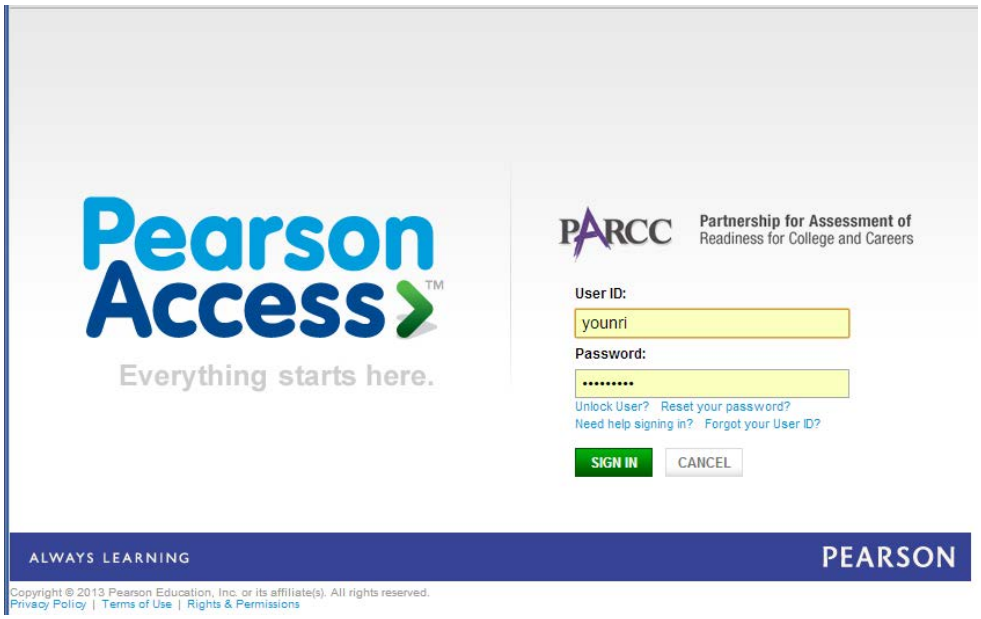
Providing Psychometric Services. Pearson psychometricians, statisticians, and research staff routinely perform and document various psychometric activities for our state assessment programs.

Comprehensive Suite of Online Assessment Services

Pearson has a long history of innovation and excellence in managing the operations of large-scale assessment programs, from original patents on Optical Mark Reader (OMR) printing and scanning to high-volume image-based performance scoring.

Pearson has also led the industry in its move to online testing and next generation assessment systems[1]. As states prepare for full-scale online testing in 2014–15, many like Arizona for the first time, Pearson has provided a wealth of resources to help ease this transition, from our online Technology Readiness Tool (TRT) used by both major assessment consortia to gauge districts' and schools' technical capacity for online testing, to onsite implementation support services, to our PearsonAccess assessment management platform.

[1] Please visit our web site at http://www.pearsonassessments.com/NextGeneration/ for full details of our research, capabilities, and experience with next generation assessments.



Single Sign-On. Authorized Arizona users will need only one login ID and password to access data applications associated with their user level in PearsonAccess.

Anticipating the industry-wide shift from paper to online testing, PearsonAccess was designed from the beginning to help state, district, and school assessment coordinators manage assessment administration and delivery simultaneosuly in both modes, with the ability to easily shift students from one mode to the other as local capacity, capabilities, and conditions change, even in the middle of the administration window. PearsonAccess also offers sophisticated user management and security profile capabilities, providing states with the ultimate in flexibility for managing and controlling access to key functionality and secure student data.

The Latest in Online Testing

Over the past decade we have implemented more than 20 online assessment programs for state and national customers, including current customers in Colorado, Florida, Georgia, Maryland, Mississippi, Tennessee, Texas, and Virginia. Other online assessments available nationally include the following:

- **The College Board's ACCUPLACER® computer adaptive testing (CAT) application** for college readiness and placement testing. Pearson delivered more than 6.3 million online ACCUPLACER assessments in 2012.
- **Our AIMSweb benchmark and progress monitoring assessments**, with more than 25 million assessments delivered each school year.
- **Pearson's NES® (National Evaluation Series),** the first large-scale teacher certification testing program aligned to professionally-accepted national learning standards and delivered entirely on the computer.

For secure, high-stakes next generation assessment delivery, our TestNav™online testing platform supports a wide range of navigation and administration options, test-taking tools, and item types, from traditional multiple-choice to open-response and highly interactive multi-media items. Built to the question and test interoperability (QTI) and accessible portable item protocol (APIP) standards, TestNav provides a

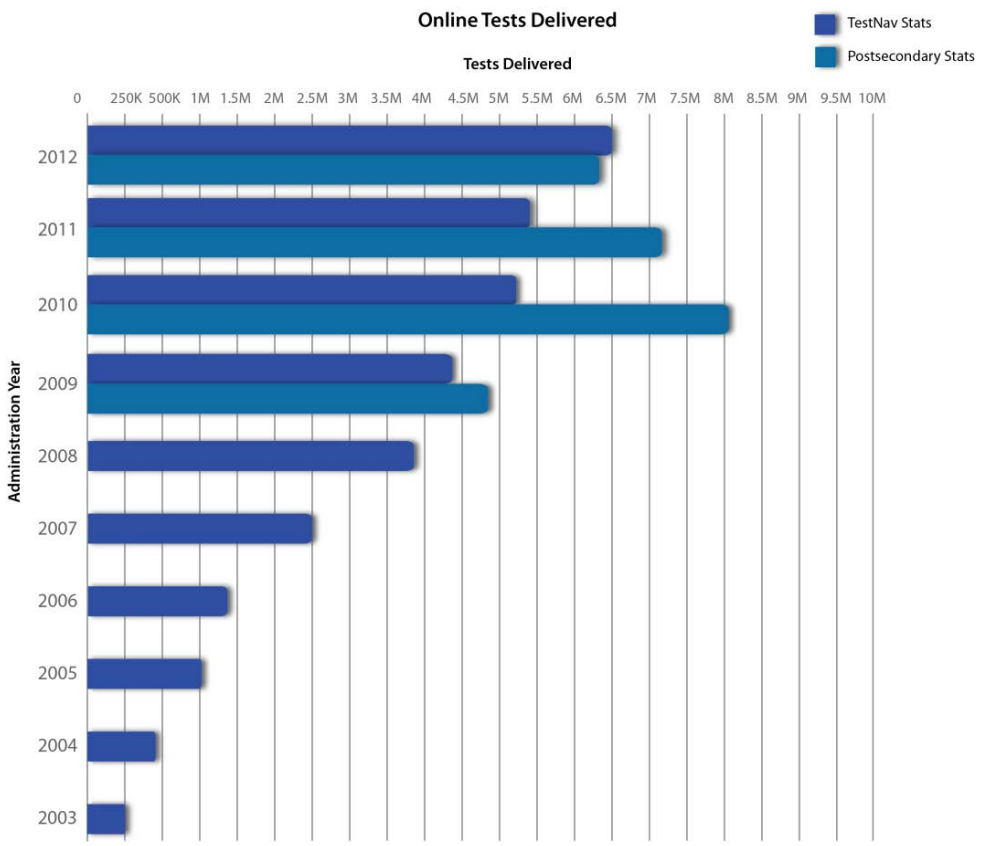
more inclusive assessment environment through numerous available online testing accommodations for students with disabilities. Now in its eighth major release, TestNav is currently used in eight states and has been selected by the PARCC consortium to deliver its spring 2014 field tests and 2015 operational assessments.

Pearson continuously invests in improving TestNav to take advantage of the latest advances in technology and to keep pace with the rapid proliferation of online K–12 instruction. At the same time, our deep roots in large-scale assessment keep us ever mindful of the stringent, high-stakes assessment requirements on which TestNav must continue to deliver:

- **Scalability** to support hundreds of thousands of simultaneous testing sessions
- **Reliability** to protect students from technical failures
- **Capacity** to deliver rich assessment content without overwhelming local internet connections
- **Compatibility** to operate on a wide range of existing devices and networks
- **Consistency** to deliver a similar experience for students, regardless of their device
- **Usability** to maximize ease of use for students, especially those with disabilities or limited technology exposure
- **Security** to protect test content and prevent student access to external resources

In addition, to meet security requirements for high-stakes testing, most online testing platforms require local technology support personnel to download and install separate software, additional components, or entirely new browsers on every testing device. For districts and schools with limited or already overburdened technology support staff, introducing online testing can present a tremendous challenge, a situation that is only compounded when updates are required. TestNav on the other hand requires no additional downloads and works with the browsers already installed on most computers, saving countless hours of technology staff time and energy that could be better spent on other important tasks.

Finally, to serve a growing number of online testing programs, we designed our system to accommodate even greater testing volumes. Because it is horizontally scalable, we can add capacity to accommodate additional students and tests. To plan for increased testing volumes during peak periods, Pearson conducts load testing in a parallel environment multiple times a year to verify that we can meet our customers' needs.



Growing Online Testing Volumes. Pearson delivered more than 12.7 million secure, high-stakes tests online in 2012.

In the first 11 months of 2013, Pearson delivered more than **11 million online tests** for 10 different assessment programs (eight state programs and two national programs). This year, we successfully delivered more than:

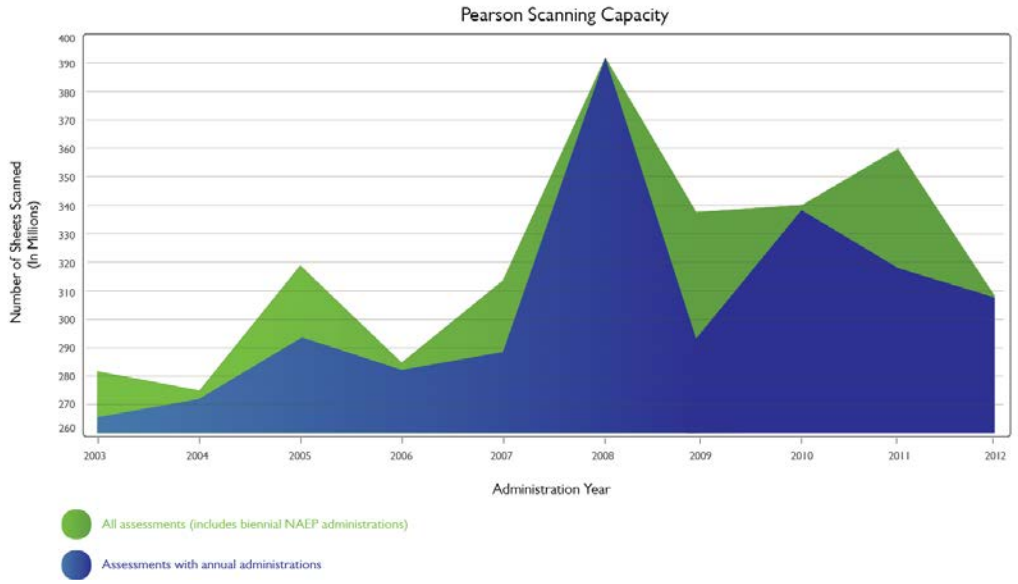
- 205,000 tests in a single hour
- 429,000 tests in a single day
- 1.2 million tests in a single week

Dependable, High-Quality Paper-Based Testing

Although the industry is moving rapidly towards online testing, a change of this magnitude does not happen overnight. Many districts and schools without sufficient capacity for online testing must continue to rely at least partially on traditional paper-based testing, while others will choose to do so for religious or cultural reasons. For these schools and their students, the need for high-quality materials, dependable service, and on-time results is equally important. Pearson's content and psychometric teams are intimately familiar with the issues associated with designing, developing, and delivering paper and online assessment that are fair, accessible, and of comparable depth and rigor.

To support our customers' ongoing needs for paper-based testing, Pearson facilities in Iowa, Minnesota, and Texas comprise more than 1.3 million square feet where we develop, design, print, package, process, scan, score, report results, and provide administrative support. Our in-house printing,

processing, and scoring operations are ISO certified, leading the industry in quality, consistency, and reliability for more than 90 years. Despite the progress that has been made in online testing, we still provide approximately 40 million paper-based assessments each year.



Pearson Scanning Capacity. Since 2000, Pearson has scanned more than 250 million sheets of student test materials each year, including 307 million in 2012.

Flexible Next Generation Scoring

Next generation assessments like ACT Aspire are taking advantage of technology to expand the use of innovative new item types that better measure the depth of knowledge and understanding inherent in the Common Core State Standards (CCSS). As the widespread use and sophistication of these assessments grows, so do concerns for cost and the ability to return scores quickly.

To meet these demands, Pearson offers a flexible array of scoring options including advanced machine-scoring algorithms, automated artificial intelligence scoring, and a database of more than 84,000 diverse, qualified scorers representing every state, more than 35,000 of whom have a master's degree or higher and more than 23,000 of whom have teaching experience.

Accurate and Efficient Automated Scoring

Machine scoring, including artificial intelligence automated scoring is a key technology supporting a next generation assessment systems. With so much riding on the successful application of automated scoring to support the practical implementation of new item/task types, having in-depth knowledge and expertise with automated scoring provides an added benefit when developing specifications for items/tasks to take advantage of this highly efficient, accurate, and reliable scoring technology.

We have recently published a number of white papers relating to the development and scoring of next generation assessments. In the white paper “Pearson’s Automated Scoring of Writing, Speaking and Mathematics” our automated scoring experts discuss the current range of automated scoring applications,

operational accuracy, and considerations for applications in next-generation assessment. This paper and others can be found at www.pearsonassessments.com/nextgeneration.

Expert Performance Scoring

Our performance scoring capabilities support a full spectrum of grade levels, subject areas, rubrics, and response types including image, text, audio, video, and portfolio scoring. In addition to scoring constructed-response items, we also offer consulting and rangefinding/benchmarking services to set appropriate scoring standards; professional development workshops for educators; and evaluation of scorability for field test item prompts.

Our scoring contracts include more than 18 state programs, as well as Washington DC and Puerto Rico programs, the National Assessment of Educational Progress (NAEP), and National Board for Professional Teachers assessments. In addition, we provided performance scoring for the 2012 ACT Aspire field tests and the 2013 ACT Aspire field studies. Furthermore, our performance scoring will be used in the operational administration of ACT Aspire in spring 2014 and fall 2014.

Arizona can benefit from our capacity and proven ability to adapt, extend, and apply technologies in new ways that advance educational assessment. Pearson was one of the first organizations to use an image-based and online distributed scoring systems to support human scoring. We went on to develop, refine, and expand these technologies to reduce paper handling, promote reliability and validity testing, and improve scoring turnaround time.

Since first scoring the NAEP in the early 1970s, Pearson has earned a reputation for stability and innovation. As an ISO-certified provider of performance scoring, our time-tested processes maintain quality standards across our scoring centers and online distributed workforce. Many of our full-time staff have teaching experience—from elementary school through post-secondary levels. Their experience and scoring capabilities extend to other state assessment content areas, including expertise in scoring English language learner and alternate assessments.

User-Friendly Reports

Pearson prints and delivers over half a million packages containing reports for more than 50 million students nationwide each year. Our production capabilities include high speed duplex, folder, and multi-page reports with full text and graphics support in color and on custom pre-printed stock.

Pearson reports aim to enhance learning at the student, classroom, school, and district levels. It is not enough to just generate reports. They must be easy to read and reflect student performance in clear and concise fashion. We design reports to enhance their usefulness for every intended audience.

Our team of report designers and systems analysts combines years of experience to deliver easy-to-interpret assessment reports. Our capabilities include paper and electronic document (PDF) reporting, media creation, online dynamic reporting, analytic reporting, and linking learning resources to student scores.

Pearson's Quality Management Systems

Our Organizational Quality Office serves as an independent advocate for product and service quality and delivery performance. This is achieved by defining and deploying effective policies that are aligned to customer needs, business strategies, and operational imperatives. Organizational Quality works collaboratively with Assessment & Instruction business units and shared services to create processes that implement these policies cost effectively and deliver desired customer and business performance results.

Domains of practice include the following:

- **Quality Assurance**, encompassing the definition, measurement, and control of product and service quality, and the continuous improvement of the processes by which we create those products and services.
- **Delivery Assurance**, encompassing the planning, measurement, and control of product and service delivery, and the continuous improvement of the processes by which we deliver.
- **Information Assurance**, encompassing the planning, measurement, and control of information protection and business recoverability.

These practices are supported by the following operating principles:

- **Assurance.** Our primary purpose is to provide quality, delivery, and information through sound policy, effective procedures, and objective measurements driving data-driven decision making.
- **Policy.** OQ policy authority encompasses Quality Assurance, Enterprise Program Management (delivery assurance), Data Security, and Business Continuity.
- **Compliance.** OQ audits policy/procedure adherence and assesses product and service conformance to customer requirements. OQ reports non-compliance and non-conformance, and confirms that corrective action is performed by process and product owners. OQ collaborates with A&I leadership to establish critical business performance indicators that trigger appropriate response actions.
- **Performance.** OQ is committed to continuously improving customer satisfaction and business capability. OQ actively and collaboratively seeks out those improvement opportunities that provide the greatest benefit to our customers and shareholders.
- **Professionalism.** OQ is the professional home of program management, quality management, security, and business continuity knowledge domains and practice disciplines in A&I. OQ leads the establishment of professional standards and professional development curricula in these domains.

The foundation of our business is a sound Quality Management System (QMS). Pearson's Assessment and Instruction (A&I) group has selected the ISO 9001:2008 standard as the guiding document for its QMS. Developed by

standards bodies throughout the world and released by the non-profit International Organization for Standardization (ISO), the ISO 9001:2008 standard offers the comprehensive quality framework necessary to demonstrate adherence to best practices and the establishment of a true culture of continuous quality improvement throughout our organization.

Pearson started on its ISO 9001 journey in 1994 and has continued to certify areas throughout the A&I organization:

- **Pearson Print Services**, Owatonna, MN, registered in 1994
- **Performance Scoring Centers (PSC),** Iowa City, IA, registered in 1999
- **Performance Scoring Centers (PSC),** Iowa City and various locations, certificate expanded to a multi-site certificate in 2001
- **Operations**, Cedar Rapids, IA, registered in 2009
- **Operations**, Austin, TX, registered in 2011
- **Operations**, Iowa City, IA, registered in 2011
- **State Services and National Services**, slated for certification in early 2014

By Our People for Your People

Customer relationships form the foundation for success—both our customers' successes and ours. We judge the success of our customer relationships by our customers' loyalty. We have provided 12 of our state customers with assessment solutions for more than 10 years, and six of those customers have trusted Pearson as their state assessment provider for **20 years or more.**

To sustain successful relationships, we hire and retain highly qualified project management professionals. We look for degreed professionals who can pay attention to details, communicate openly, and demonstrate the drive to complete tasks and activities required for on-time, accurate delivery. We support our staff with regular training opportunities so they stay knowledgeable and informed, which improves their responsiveness and performance to our customers.

Our customers reap the direct benefits of the breadth and depth of experience in the management and support teams we provide and the level of authorization each team has to meet the requirements of the specific assessment project.

Certified Program Management

Our customers benefit from our coordinated, professional approach to project management. This we deliver through Pearson program teams that follow proven management techniques identified by the Project Management Institute (PMI®), an international non-profit organization that provides a consistent management framework, the Project Management Body of Knowledge (PMBOK).

As an additional demonstration of our attention to quality program management, Pearson program managers also possess, or are actively pursuing, PMI Project Management Professional (PMP) certification. The PMP certification program objectively assesses and measures professional knowledge.

Our program managers bring a variety of experience to their positions. Some have classroom or education administration experience, while others have worked for years in different areas of Pearson or for other testing companies. All of this experience brings critical management, education, and testing knowledge to the programs they serve.

A strong project management plan defines how a project will be executed, monitored, and controlled. Arizona's assessment systems have benefitted from our adherence to the PMI management model, which promotes quality and on-time delivery throughout a project's duration. The PMI framework encourages our program team to determine program requirements, adhere to a detailed project plan, effectively manage scope changes, and follow sound business practices to minimize risk. In turn, this helps us achieve greater predictability and repeatability of results.

Efficient, Responsive Customer Service

The service and responsiveness testing contractors provide to local assessment coordinators, technology personnel, and other constituents is often viewed as an extension of the service provided by their own state department of education or sponsoring organization. As such, we take seriously our responsibility to uphold our the reputation of our customers' programs when assisting those who have questions or need additional assistance.

Our Customer Service Center (CSC) uses a combination of online and in-house knowledge base and call routing systems within a team-centered approach to deliver excellent service. Using our quality processes and best practices, our CSC serviced almost 105,000 contacts in 2012. Customers surveyed indicated that their expectations were met or were exceeded 93 percent of the time.

When local test coordinators call for customer support, they want to talk to a live person who is a qualified professional and able to offer immediate assistance. Pearson provides experienced, competent customer support staff who can answer questions efficiently and thoroughly.

To deliver on that promise, we require that our CSC staff demonstrate excellent customer service skills in addition to appreciating that each caller is an individual with specific needs and concerns. We ask that new staff have a minimum of six months experience in customer service, help desk, or telemarketing, including prior experience identifying strengths or weaknesses of alternatives and recommending appropriate decisions and solutions.

Responsibility to Care and to Give

For us, "education" encompasses reaching out to others, and making a difference every day. Embodying the role of good corporate citizen is part of our DNA, characterized by our three core values: brave, imaginative, and decent.

Our parent corporation, Pearson plc, combines a commitment to its goals with a clear understanding of its responsibilities in the wider world. These wider responsibilities include charitable giving focused on education and literacy, involvement in education-related associations and organizations, and commitment to sustainable practices concerning society and the environment.

The Pearson Foundation

The philanthropic arm of Pearson plc, The Pearson Foundation, extends Pearson's commitment to education by partnering with leading nonprofit, civic, and business organizations to provide financial, organizational, and publishing assistance across the globe. The Foundation aims to make a difference by sponsoring innovative educational programs and extending its educational expertise to help in classrooms and in local communities. Key initiatives include the following:

- Read For The Record, a US-based celebration of reading and learning that each year aims to set a new world record for the most children reading the same book on the same day. On Oct. 7th, 2010, 2,057,513 people across the globe did just that, reading the same classic children's book, The Snowy Day by Ezra Jack Keats.
- **The Pearson Affordable Learning Fund** makes minority equity investments in for-profit companies that help meet a burgeoning demand for affordable education services in Africa, Asia, and Latin America.
- The Pearson Prize for Higher Education, a two-tiered award that celebrates US-based students that give back to their college communities and helps them continue their work.
- The Pearson Fellowship for Social Innovation, which supports exemplary young leaders around the globe who are using their ingenuity, passion, and energy to build better lives for themselves and their communities.

Caring for Our Environment

Environmental responsibility is an essential component of what it means to be a good corporate citizen, and protection of the environment is embedded in everything that Pearson does. One hundred percent of the electricity used by Pearson in the US comes from renewable sources; we purchase through local power utilities and purchase credible, third-party verified renewable energy credits.

Pearson was the first global media company to commit to becoming carbon neutral, a goal we achieved in 2009. Since then, our businesses have been working diligently to continue to reduce their environmental impacts. Our Old Tappan, NJ, on-demand print center is Pearson's first site to be powered directly by solar energy. Over the 25-year life of the panels, the switch to solar energy is expected to offset about 4,000 tons of carbon dioxide, the equivalent of planting more than 10,000 trees. Our first wind turbine, built at Pearson's Print Services location in Owatonna, MN, will offset approximately 119 metric tons of carbon dioxide annually, or 2,830 metric tons—the equivalent of 5,540 barrels of oil—over its 20-year life. And the solar panel installation at our Cranbury, NJ facility will be one of the largest single site installations of solar panels in the world.

Pearson in the US was named as a 2010 Green Power Leadership Award winner by the U.S. Environmental Protection Agency (EPA). Pearson currently ranks No. 25 on EPA's National Top 50 list of green power purchasers.

Our Financial Stability

Pearson's continued stability, even in perilous economic times, enables us to invest in our company and to conduct research, innovate, and continually improve our processes and services.

The last 10 years of financial reports are available at www.pearson.com/investors/financial-information. Adjusted earnings per share, dividends per share, and total free cash flow per share are provided in pence, and return on invested capital is provided as a percentage; the remaining figures are £ millions.

ACT Aspire: Born of Two Longtime Collaborators

Arizona's transition to new grade-level assessments in spring 2015 can benefit from next-generation assessments offered by ACT Aspire, LLC, a new company born of a half-century collaboration between parent organizations ACT, Inc. and Pearson.

Created to combine the recognized expertise and capabilities of those co-parents, ACT Aspire, LLC, delivers exceptional assessment solutions that can inform Arizona students, parents, educators, and other stakeholders where individual students stand on the path to career- and college-readiness, and where they are headed.

ACT Aspire, LLC, licenses assessments from both ACT, Inc. and Pearson through a defined, joint-venture agreement. With Pearson serving as prime contractor to Arizona, and ACT Aspire, LLC, as a subcontractor, each organization can contribute resources as needed to provide successful implementation, management, and delivery of next generation assessments like the proposed Arizona SBCA.

ACT Corporate Capabilities

ACT programs and services for educational assessment, workforce development, career planning, professional certification and licensure, data management, and research have become mainstays of the American educational and business landscapes. By continuing to invest in superior staff, technology, and resources, ACT remains a leader in providing quality assessment and related services.

Leading in Education

In the 1950s, ACT's founders recognized that the United States was being transformed from a society in which only the elite attended college to one in which the majority would continue their education beyond high school. Inundated with applications from in-state students with widely disparate levels of preparation and from out-of-state students who attended unfamiliar high schools, colleges sought more and better information on which to base their admission and placement decisions.

In 1958, E. F. Lindquist, University of Iowa professor, psychometrician, and inventor of the first optical-mark scanner, defined his goals for the college entrance examination that would be ACT's genesis—the

American College Testing Program. The program would test broad competencies, rather than rote memorization, and would encourage students to acquire knowledge and learn to use it in creative ways.

Committed to helping students present themselves to colleges as unique individuals—not just as test scores—ACT has over the years expanded and refined the information we collect and report. Today, in addition to scores in English, math, reading, and science, the ACT provides details about students' interests, needs, plans, and goals. These resources are used for admission, recruitment, course placement, advising, and counseling. This year—for the second consecutive year—more graduates in the class of 2013 took The ACT to demonstrate their college and career readiness than took the SAT (nearly 1.8 million more).

The ACT Aspire Assessment System

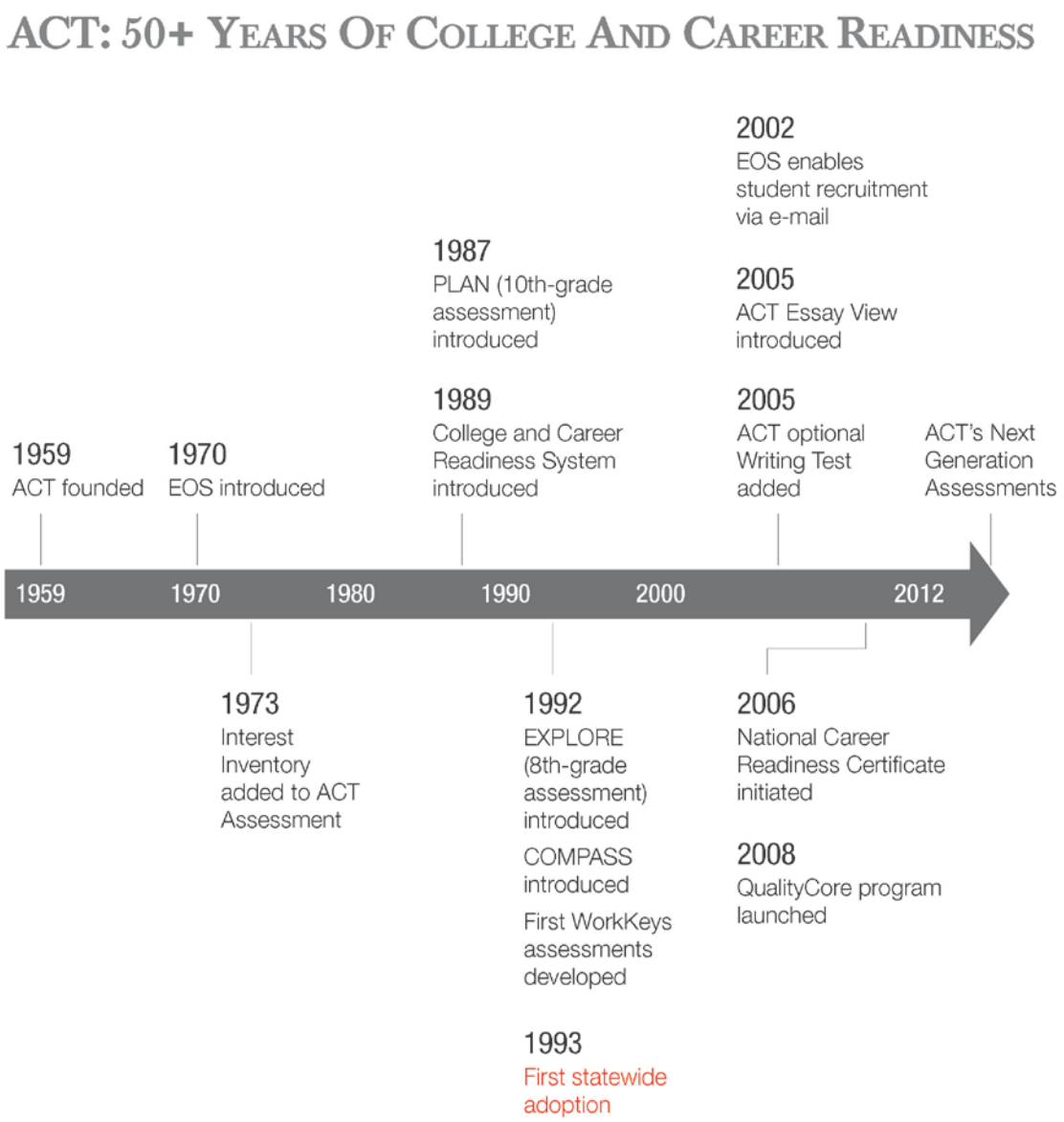
ACT is preparing for a spring 2014 launch of the ACT Aspire Assessment System® (ACT Aspire®) in response to the longstanding need for an integrated, longitudinal assessment system that addresses the gaps between the skills students are learning in school and the skills they need to succeed in college and careers in an increasingly competitive global economy.

ACT Aspire is a computer-based, longitudinal assessment system that connects growth and progress from elementary grades through high school in the context of college and career readiness. Assessing student knowledge and skills is critical to providing students with the support necessary to help them succeed in school and life. The essential, actionable information and insights gained will help educators understand a student's past, capture the present, and affect the future.

The new longitudinal assessment system—anchored by the ACT—will offer an integrated, multidimensional approach to college and career readiness that focuses on measuring achievements and behavior relative to goals. It will be the first digital, longitudinal assessment system to fully connect student performance from early elementary to high school, helping students know exactly where they are and providing insights on how to build on strengths and address weaknesses, both in and out of the classroom. The system will be linked to ACT's College Readiness Standards™ and Benchmarks.

Common Core Aligned

ACT Aspire is aligned with the Common Core State Standards and Arizona's College and Career Ready Standards, which draw heavily from ACT's College Readiness Standards and research. It measures English, mathematics, and reading skills and provide insight on science skills and writing skills from grades 3–10. In addition, in grades 9–10, ACT Aspire will provide a Work-Ready Score and a predicted ACT Composite Score.



Leading in Education. ACT has extensive experience developing and delivering large-scale, high-stakes K–12 assessment programs designed to determine college and career readiness.

Working with Pearson to Serve Arizona

ACT's leadership in educational assessment combined with the technology expertise of Pearson makes this new system a strong match for Arizona as the state strives to prepare more students for college and careers in a changing world.

[redacted]

[redacted]

Working Together: Arizona and ACT

Arizona ACT Initiative

Strongly committed to improving student college and career readiness, Arizonans have worked with ACT in multiple ways to promote this goal. Since 2009, several schools have participated in the Arizona ACT Initiative, which has significantly increased the number and diversity of students in the state completing The ACT. In spring 2013, more than 19,000 public school juniors from 14 districts participated. Due in part to this initiative, among 2013 graduates who completed the ACT, more students identified themselves as Hispanic/Latino than any other ethnic group for the first time in Arizona history.

GEAR UP

Northern Arizona University GEAR UP has promoted and underwritten the costs associated with completing the ACT Explore Assessment for grade 8 students at schools with 50 percent or more free and reduced lunch students. As reported on the GEAR UP Website, “The 2012–13 Explore program had 269 schools representing 37,427 eighth graders participating, resulting in 38 perfect scores.” The goal of this program is to use individual and aggregate grade 8 college and career readiness student data to improve both the student’s ability to matriculate and succeed in a postsecondary environment and the school system’s ability to create more and better prepared high school graduates.

Move on When Ready

ACT QualityCore serves as one of two aligned instructional systems certified by the National Center on Education and the Economy and approved by the Arizona State Board of Education for use in both the lower (grade 9 and 10) and upper (grade 11 and 12) divisions as part of Arizona’s Move on When Ready and Grand Canyon Diploma programs. In partnership with the Center for the Future of Arizona, with funding from the Helios Education Foundation, ACT supported five partner schools’ successful implementation and delivery of the ACT QualityCore program in 2012–13.

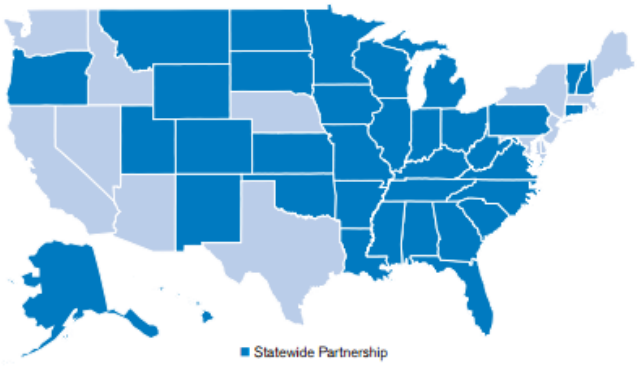
Achieving Educational and Workplace Success

The solutions we offer in response to your RFI are designed to build upon Arizona’s current initiatives and demonstrated successes. ACT has a history of more than 50 years in delivering sustainable innovation born of high quality research and empirical education data. Though designed to meet a wide array of needs, ACT programs and services have one guiding purpose—to provide information to help people achieve educational and workplace success.

Resources

Statewide Partnerships in College and Career Readiness

States that incorporate ACT's college and career readiness solutions as part of their statewide assessments provide greater access to higher education and increase the likelihood of student success in postsecondary education. Educators also have the ability to establish a longitudinal plan using ACT's assessments, which provide high schools, districts, and states with unique student-level data that can be used for effective student intervention plans.



State administration of ACT's programs and services:

- Increases opportunities for minority and middle- to low-income students.
- Promotes student educational and career planning.
- Reduces the need for remediation.
- Correlates with increases in college enrollment, persistence, and student success.
- Aligns with state standards.

ACT Aspire 3rd- through 8th-grade students	ACT Explore 8th- and 9th-grade students	ACT Plan 10th-grade students	™ACT 11th- and 12th-grade students	ACT QualityCore 8th- through 12th-grade students	ACT WorkKeys 11th- and 12th-grade students	ACT National Career Readiness Certificate™	ACT WorkKeys®-based certificates
Alabama	Alabama	Alabama	Arkansas	Alabama	Alaska	Alaska	Alabama
	Arkansas	Arkansas	Colorado	Kentucky	Illinois	Connecticut	Arkansas
	Hawaii	Florida	Hawaii		Hawaii	Indiana	Colorado
	Illinois	Hawaii	Illinois		Michigan	Iowa	Florida
	Kentucky	Illinois	Kentucky		North Carolina	Kentucky	Georgia
	Louisiana	Kentucky	Louisiana		North Dakota	Louisiana	Indiana
	Michigan	Louisiana	Michigan		Wyoming	Michigan	Kansas
	Minnesota	Michigan	Montana			Minnesota	Mississippi
	North Carolina	Minnesota	North Carolina			Missouri	North Carolina
	Oklahoma	New Mexico	North Dakota			Montana	Oklahoma
	South Carolina	North Carolina	Tennessee			New Hampshire	South Carolina
	Tennessee	Oklahoma	Utah			New Mexico	Virginia
	Utah	Tennessee	Wyoming			North Dakota	West Virginia
	West Virginia	Utah				Ohio	Wyoming
	Wyoming	West Virginia				Oregon	
		Wyoming				Pennsylvania	
						South Dakota	
						Tennessee	
						Vermont	
						Wisconsin	

Statewide Partners. ACT programs and services help states increase student opportunities, promote educational and career planning, reduce the need for remediation, and increase college enrollment and student success.

Requirement

b. Identify the individuals from the organization that will be working with Arizona officials on all aspects of the assessments' implementation.

Response

Providing Arizona a Talented Program Team

A cohesive, well-rounded team is vital to building a strong project plan, executing the plan, and building trust with customers and constituents. Arizona will continue to benefit from the emphasis we place on our people and the teams assembled to deliver services to our customers.

Rich Young, State Services Vice President, will serve as the primary point of contact for Arizona officials on all aspects of the ADE's assessment information-gathering and procurement processes. With more than five years of direct management and support for Arizona's assessment systems and nearly a decade of involvement with Arizona's education system overall, Mr. Young maintains a broad network of contacts throughout Arizona's education community and is deeply familiar with Arizona's reform goals, objectives, and initiatives. He will serve as a conduit to senior management and executives at both Pearson and ACT to help Arizona stakeholders further explore any specific areas of interest.

Should the ADE ultimately choose to implement one of the assessment options presented in our response, whether Pearson serves as prime contractor or subcontractor, Mr. Young and his current Arizona team will also be prepared to help support your transition to new assessments in whatever way possible, and we will make every effort to see that we remain involved with your programs to provide smooth continuity for as long as possible. If you should have any questions or need further information, please feel free to contact Mr. Young at 319.621.4042 or via email at rich.young@pearson.com.

2. Overview of Assessment

Requirement

a. Describe the assessments, the grades assessed, the subject areas included, and the formative, diagnostic and summative components. Also describe available end-of-course assessments for grades nine through eleven.

Response

For Arizona's Standards-Based Competency Assessments Grades 3–11 (SBCA), Pearson can provide two high-quality options, each with their own unique benefits, each of which can be delivered through Pearson's proven online and paper-based assessment systems:

1. Existing assessments aligned to ACT's College Readiness Standards™ and Benchmarks, which served as one of the foundations upon which the Common Core State Standards (CCSS) and Arizona's College and Career Ready Standards were created.
2. A custom assessment built to a new test blueprint of Arizona's own design, from a combination of Arizona-owned and licensed third party content aligned to Arizona's College and Career Ready Standards

Should the ADE elect to implement a third option through PARCC, Pearson would continue to serve Arizona indirectly as PARCC's online test delivery subcontractor through the spring 2015 operational assessment. Pearson is strongly committed to PARCC's success and we continue to pursue opportunities that would expand our service to PARCC beyond online delivery in spring 2015. However, due to the ADE's potential role in PARCC procurements, it would be inappropriate for us to further elaborate on the PARCC assessments or speculate on the potential scope of Pearson's future support for PARCC in this response.

Measuring Growth to Readiness through ACT's College Readiness Standards and Benchmarks

Pearson, ACT, and ACT Aspire can provide a comprehensive solution featuring a variety of components that can work together to provide a full picture of students' progress toward college and career readiness. For grades 3–10, we offer ACT Aspire™, a computer-based, longitudinal assessment system that connects growth and progress from elementary grades through high school in the context of college and career readiness.

For end-of-course assessments, we offer ACT QualityCore®. ACT QualityCore end-of-course assessments are built to measure mastery of rigorous course standards for core high school courses.

The capstone of our solution for grade 11 would be the well-known ACT™ Assessment used to assess students' academic readiness for college.

In this section we have provided an overview of each potential component that could be used to craft a solution for Arizona.

ACT Aspire

Putting Arizona Students on Track for Success

For the SBCA, Pearson and ACT, Inc. offer ACT Aspire™ for grades 3 to 10, accompanied by the nationally recognized ACT™ assessment for grade 11.

ACT Aspire represents our best solution to help students get on target and stay on target for college and career readiness. The capstone ACT college entry and placement exam provides an additional, widely accepted measure of college readiness and helps Arizona high school seniors eliminate at least one barrier to postsecondary education: the college entrance exam.

Aligned from elementary grades to high school, connecting each grade level to the next, ACT Aspire offers Arizona a cohesive longitudinal system that enables teachers and parents to know with confidence, beginning at grade 3, where students are on the path to college and career readiness.

Product of a Long Collaboration

ACT Aspire is the latest program born of a 50-year collaboration between parent companies ACT and Pearson. Both companies have successfully developed advanced assessment and learning projects for more than half a century—a distinct advantage of ACT Aspire

Offering an End-to-End Solution for Arizona

With ACT Aspire, Arizona will get an integrated, multidimensional approach to college and career readiness that focuses on measuring achievements relative to clearly-defined goals. ACT Aspire is the first digital, longitudinal assessment system to fully connect student performance from early elementary to high school, helping students know exactly where they are while providing insights on how to build on strengths and address areas in need of improvement. ACT Aspire offers more subjects, deeper research, and proven technology to Arizona educators.

More Subjects

ACT Aspire offers Arizona the following benefits:

- Assesses student progress in five critical subject areas—math, reading, English, writing, and science
- Links to ACT's validated College Readiness Standards™ and College Readiness Benchmarks™
- Aligns to the Arizona College and Career Ready Standards, which borrow heavily from ACT's College Readiness Standards and research

In answer to Arizona's desire for competency assessments aligned to Arizona standards, our solution meets the spectrum of ADE requirements.

The ACT Aspire package includes summative, interim, and classroom assessments. We offer proven program management, a web-based enrollment and order collection system to ease the burden on Arizona school staff, an online assessment system to accommodate the vast differences in size and Internet connectivity in Arizona schools, and a carefully thought-out transition from paper and pencil to online assessment.

Meeting Arizona's Needs

To meet Arizona's needs for transition to online assessments as well as assessments aligned to the Arizona standards, ACT Aspire offers the following:

- Computer-based test administration with paper and pencil as an option and also for accommodations
- Inclusion of multiple item types—selected response, constructed response, performance events and technology enhanced
- Online reporting for administrators, students and parents
- A direct writing assessment in addition to English, math, reading, and science
- Vertical scale alignment from grade 3 through high school
- Insights on additional dimensions of readiness (e.g., text complexity, justification and explanation)
- Alignment to an interim/classroom-based assessment tool

Informing Progress Toward College and Career

For the critical advantage of forecasting success with college entry exams, ACT Aspire offers a particularly distinguishing feature—a trajectory from grade 3 all the way to the renowned ACT college readiness assessment and its familiar 1 to 36 score scale. In grades 3 through 7, ACT Aspire provides a three-digit score and in grades 8 through 10, the ACT Aspire provides a predictor of student performance on The ACT. From year to year, the scores are vertically linked, essentially offering a view of preparedness from third grade all the way to college.

Actionable insights comprise the central defining features of ACT Aspire. With these insights, educators and parents will know if a student is on track for college and career through data garnered from either the Arizona College and Career Ready Standards or the ACT College and Career Ready Standards. However, among current assessment systems, only ACT Aspire empowers educators to draw insights from both.

Measuring Areas Critical to Success. ACT Aspire measures the following critical areas needed for student success:

- English, writing, reading, math and science
- Text complexity
- Science, technology, engineering, and mathematics (STEM) and Work Readiness

ACT Aspire employs modular administration by subject and/or grade and provides evidence-based results.

ACT Aspire score reports will focus on students' journeys through the following guiding principles:

- Progression toward **readiness** matters, not just the individual scores
- Score reports will tell a **story**, reflecting on each child's possibilities, not his or her deficits
- The reports will have **easy to understand language** for all audiences

ACT Aspire's comprehensive reports will provide scaled subject scores in English, math, reading, science, and direct writing. **Progress charts** will demonstrate student longitudinal growth against The ACT Readiness Benchmark for each subject and a two-year student "predicted path" toward college and career readiness. Our solution provides **national norms** as well as ACT Aspire Composite Scores and Predicted ACT Subject/Composite Scores (grades 9 and 10 only).

ACT Aspire Interim and Classroom Assessments

Also available to support the ACT Aspire Summative assessments for grades 3–10 are ACT Aspire interim and classroom assessments, designed to generate timely and actionable data that may be used to provide progress and growth measures along with a snapshot (comparison) as defined by the end user. The ACT Aspire interim and classroom assessments have the following three goals in mind:

- To give teachers information to make instructional decisions
- To give students information about how to improve their academic performance
- To provide students and educators with individualized pathways to improve

Available in the 2014–15 academic year, the ACT Aspire interim and classroom assessments will include multiple-choice items only (constructed-response and technology-enhanced items will be added as research allows), be based on the ACT Readiness Standards, aligned to the CCSS and Arizona's College and Career Ready Standards, and include report data usable for item response/analysis. [redacted]

Throughout the 2014–15 academic year, ACT Aspire interim assessments will include three benchmark assessments delivered in grades 3–10, as follows:

- The Interim I baseline assessment is designed to provide educators with a snapshot indicating where students' knowledge is at the beginning of the academic year.
- The Interim II assessment is designed to provide educators with information to make instructional decisions.
- The Interim III assessment is designed to allow educators an opportunity to assess what they have taught based on their local curriculum pacing/alignment, along with simulating the summative assessment experience as it relates to amount of content covered at this point during the academic year.

[redacted]

ACT Aspire interim test administration times are drafted to be in line with summative (perhaps slightly shorter), and are predictive of how a student is likely to perform toward meeting a benchmark. The interim assessments will have the same test development processes and data collection as the summative assessments. Interim assessments can be used to make data-based inferences about examinees' college and career readiness. ACT Aspire interim assessments are curriculum-neutral, and once they are empirically linked to the ACT Aspire summative assessments, can be used to infer and measure student growth.

The ACT Aspire Classroom assessments solution also will provide the following:

- Classroom assessments that align to ACT College Readiness Standards
- Classroom assessments that align to the CCSS and Arizona's College and Career Ready Standards
- Classroom assessments that inform daily instruction
- Classroom assessments that include a range of metacognitive items that help students increase their awareness of their own cognitive processes (i.e., self-assessment)

[redacted]

- Assessment results provided as immediate feedback with links to instructional resources

[redacted]

For deeper insight, visit www.discoveractaspire.org and learn more about the value of ACT Aspire. This is a one-of-a-kind initiative, borne of partner parents with strong experience in content, research, and technology. We stand ready to support Arizona students, parents, educators, and other stakeholders on the path toward career and college readiness for every student.

ACT QualityCore

The ACT QualityCore program reflects the attributes of instructional systems in high-performing countries through a syllabus-driven, integrated system of instruction that includes instructional materials, aligned formative items and summative assessments, professional development, and an unwavering focus on developing the critical thinking and analytical knowledge and skills students need when they leave high school. Given these attributes, ACT QualityCore is well positioned to support the implementation of comprehensive end-of-course assessments for high schools in Arizona.

ACT QualityCore is Research-Based

ACT QualityCore is a research-based instructional improvement system of end-of-course assessments, educator's resources, formative items, and benchmark assessments designed to help schools better prepare students for college and careers.

ACT QualityCore is based on course standards that are rigorous, empirically-based and derived from the course offerings at high-performing high schools in the United States. The goals for the program are to:

- Improve student achievement
- Ensure that course content is focused on college and workforce readiness standards as well as state standards
- Provide schools and school districts with assessment data about student progress to help improve teaching and learning

ACT QualityCore is structured so that high school core courses are focused on the most essential postsecondary skills and enables educators to monitor progress toward and attainment of those skills. The ACT QualityCore subjects included in this response are:

- English 9, English 10, English 11
- Algebra I, Geometry, Algebra II

Redefining Rigorous Content for High School Core Courses

Through an intensive multiyear study of classroom practice, ACT's research resulted in core course syllabi that were tightly aligned with postsecondary expectations and drove the development of curriculum and instruction, formative and summative assessments, and professional development for teachers and leaders. This research also played a major role in the development of the Common Core State Standards (CCSS), resulting in a tight alignment between the CCSS and ACT QualityCore course standards.

ACT QualityCore has been developed to:

- Improve students' college and career readiness so that they are ready to enter two-year, four-year, trade or technical schools or workforce training programs after high school without needing remediation in any subject area

- Confirm that each course is a highly integrated instruction/learning system that is focused on the essential standards that students need to learn to be ready for college and career (i.e., the Arizona College and Career Ready Standards)
- Provide schools and districts with course-level data to improve teaching and learning by examining the value added by each course to college and career readiness through course quality and impact

ACT QualityCore also serves as one of two aligned instructional systems certified by the National Center on Education and the Economy (NCEE) and approved by the Arizona State Board of Education for use in both the lower (grades 9 and 10) and upper (grades 11 and 12) divisions as part of Arizona's Move on When Ready and Grand Canyon Diploma programs. In partnership with the Center for the Future of Arizona, with funding from the Helios Education Foundation, ACT supported five partner schools' successful implementation and delivery of the ACT QualityCore program in 2012–13.

The ACT

A Superior Solution: The ACT in 11th Grade

Arizona will benefit from an assessment system that addresses the gap between the skills students are learning in school, and the skills they will need to succeed in college and careers in the increasingly competitive global economy.

This section specifically describes the ACT In-School program we propose as the grade 11 end-of-high school assessment for the SBCA. While ACT Aspire grades 3-10 helps students remain on target for college and career readiness, the capstone ACT Assessment serves both as the final measure of college readiness, and is accepted by all four-year colleges and universities in the United States.

Note: While science was not included within the scope of this RFI and therefore was specifically excluded from the ACT Aspire 3–10 assessments (though it is available as a separately priced option), science is included in The ACT. English, mathematics, reading, science, and writing are skills that are most important for success in postsecondary education and that are acquired in secondary education. The ACT is designed to measure all these skills for college entry and placement.

A Strategic Partnership Spanning 50 Years

For more than 50 years, Pearson and ACT have worked together to support complex, large-scale assessment projects both in K–12 and the college and career readiness market segments. Since 1960, Pearson has provided test processing and scoring services for the ACT college admissions and placement test taken by more than 1.8 million high school graduates each year. In 13 states, including Michigan, Kentucky, Colorado, Wyoming, and North Dakota, the ACT is administered as part of a state testing program using the same in-school testing model we propose for Arizona. Currently, 15 state programs use ACT's Explore (8th grade) and 16 states administer ACT's Plan (10th grade) on a statewide basis.

ACT Test Design

One fundamental idea underlies the development and use of ACT's assessments: The best way to determine how well prepared students are for further education and for work is to measure as directly as possible the knowledge and skills needed in those settings in ways that can be supported by evidence. Accordingly, curriculum study in the content areas (English, mathematics, reading, and science) is ongoing at ACT and includes reviews of tests, curriculum guides, and national standards; a one-of-a-kind nationwide survey of educational practices and expectations; and meetings with content experts. Information and evidence from these sources defines the scope and sequence for the areas measured by ACT assessments.

The ACT National Curriculum Survey®, conducted by ACT every three to four years, continues to play a primary role in guiding the development of ACT's assessments. ACT surveys thousands of elementary school, middle/junior high school, secondary, and postsecondary teachers in English/writing, reading (including English language arts and social studies teachers), mathematics, and science to determine what skills and knowledge are currently being taught and which are considered important for success at each grade level for college and career readiness. The survey data ACT collects about what entering college students need to know and be able to do to be ready for college-level coursework helps inform ACT's assessments and confirm that they meet the needs of college and career readiness.

The ACT English, Mathematics, Reading, Science, and Writing Tests

The ACT contains four multiple-choice tests—English, mathematics, reading, and science—and a writing test. These tests are designed to measure skills that are most important for success in postsecondary education and that are acquired in secondary education. [redacted]

The ACT tests are oriented toward the general content areas of college and high school instructional programs. The test questions require students to integrate the knowledge and skills they possess in major curriculum areas with the information provided by the test. Thus, scores on the tests have a direct relationship to the students' educational progress in curriculum-related areas and possess a meaning that is readily grasped by students, parents, and educators.

The ACT English Test is a 75-item test that measures understanding of the conventions of standard written English (punctuation, grammar and usage, and sentence structure) and of rhetorical skills (strategy, organization, and style). Spelling, vocabulary, and rote recall of rules of grammar are not tested. The test consists of five prose passages, each accompanied by a sequence of multiple-choice test items. Different passage types are employed to provide a variety of rhetorical situations. Passages are chosen not only for their appropriateness in assessing writing skills, but also to reflect students' interests and experiences.

The ACT Mathematics Test is a 60-item test designed to assess the mathematical reasoning skills that students across the United States have typically acquired in courses taken up to the beginning of grade 12. The test presents multiple-choice items that require students to use their mathematical reasoning

skills to solve practical problems in mathematics. Knowledge of basic formulas and computational skills are assumed as background for the problems, but memorization of complex formulas and extensive computation are not required. The material covered on the test emphasizes the major content areas that are prerequisite to successful performance in entry-level courses in college mathematics. Six content areas are included: pre-algebra, elementary algebra, intermediate algebra, coordinate geometry, plane geometry, and trigonometry.

The ACT Reading Test is a 40-item test that measures reading comprehension as a product of skill in referring and reasoning. That is, the test items require students to derive meaning from several texts by referring to what is explicitly stated and reasoning to determine implicit meanings. Specifically, items ask students to use referring and reasoning skills to determine main ideas; locate and interpret significant details; understand sequences of events; make comparisons; comprehend cause-effect relationships; determine the meaning of context-dependent words, phrases, and statements; draw generalizations; and analyze the author's or narrator's voice or method.

The ACT Science Test is a 40-item test that measures the interpretation, analysis, evaluation, reasoning, and problem-solving skills required in the natural sciences. The content of the Science Test is drawn from biology, chemistry, physics, and the Earth/space sciences, all of which are represented in the test. Students are assumed to have a minimum of two years of introductory science, which ACT's National Curriculum Studies have identified as typically one year of biology and one year of physical science and/or Earth science. Thus, it is expected that students have acquired the introductory content of biology, physical science, and Earth science, are familiar with the nature of scientific inquiry, and have been exposed to laboratory investigation.

The ACT Writing Test is an essay test that measures students' writing skills—specifically those writing skills emphasized in high school English classes and in entry-level college composition courses. The test consists of one writing prompt that defines an issue and describes two points of view on that issue. The students are asked to respond to a question about their position on the issue described in the writing prompt. In doing so, they may adopt one of the perspectives described in the prompt, or they may present a different point of view on the issue. The essay score is not affected by the point of view taken on the issue.

[redacted]

Custom Assessment Option

If Arizona should determine that a shared or pre-built assessment will not fully meet the state's needs, the ADE could procure a custom assessment built to a blueprint of Arizona's own design, using a combination of high-quality, well-aligned AIMS items and items licensed from Pearson and/or one or both of the national assessment consortia. The potential benefits of this approach include the following:

- Blueprints could be created to reduce total testing time and/or the frequency of items that require expensive hand scoring without significantly impacting overall assessment quality.
- Arizona would have greater freedom to specify accommodations and create custom test specifications for parallel paper-based testing, if desired or as needed.
- Items written for the PARCC and SMARTER Balanced Assessment Consortium (SBAC) would provide strong assessment validity, due to their established alignment with the Common Core and Arizona's College and Career Ready Standards.
- These pools would also include a wide variety of challenging item types, including machine-scorable selected-response and technology-enhanced items, those that can be scored automatically with artificial intelligence, and constructed-response items that require human scoring.
- Recently-developed AIMS items have been written to greater depths of knowledge (DOK) and rigor and have been thoroughly vetted by the ADE and Arizona educators. Many or most of these items would also be appropriate for use in Arizona's SBCA.
- Because AIMS items are owned by Arizona, they can also help reduce cost by displacing third-party licensed content.
- Most AIMS items and any additional items licensed from PARCC for use in 2015 operational assessments would have already been field tested with a large sample of Arizona students, resulting in stronger operational reliability and validity. If needed, additional items not previously tested in Arizona could be embedded for field testing in spring 2015 to build the available pool for subsequent years' assessments.
- As a result of our role as the ADE's current contractor and the PARCC 2014 field test contractor, Pearson would already have many of these items in our item banking systems, reducing transition cost and lead time for building the 2015 operational forms.
- Pearson is already familiar with the ADE's established processes for creating test blueprints and selecting from pools of established items to build high-quality operational assessments, again helping to streamline the transition to a successful 2015 operational startup.
- If sufficient items are available for use, interim, benchmark, and/or diagnostic assessments could also be designed for use at key points or throughout the school year.
- Pearson will also have detailed knowledge and established processes for scoring all or most AIMS and PARCC items used, helping to further reduce startup cost and minimize implementation risk.
- It would likely be possible to maintain national score comparisons by using PARCC items as a link for equating to the full PARCC assessments, similar to the way Arizona has used Stanford 10 items for national comparisons in the past.

The only major drawback to maintaining a custom assessment is that Arizona could experience diminished ownership as existing AIMS items are retired from use over time, or be faced with additional cost to develop new items (depending on PARCC's future development and the nature and extent of customization in Arizona's SBCA).

Requirement

b. Describe the timeline for the development of the assessments to ensure full implementation by the 2014 –2015 academic year.

Response

ACT Aspire

The ACT Aspire assessments began development in 2012 and will be operational in April 2014. Field tests and studies were conducted in late 2012 and spring 2013 to establish the operational forms. Multiple forms will be built to support future ACT Aspire administrations (i.e., the 2014-2015 school year). Additionally, the forms development cycle will continue throughout 2014 and beyond to maintain the necessary forms for test administrations beyond the 2014-2015 school year.

ACT QualityCore

ACT QualityCore is a fully developed, fully functioning research-based instructional improvements system of end-of-course assessments, educator's resources, formative items, and benchmark assessments designed to help schools better prepare students for college and career. Approximately 300,000 ACT QualityCore tests were taken in 2012.

The ACT

The ACT College Readiness Assessment is well known and valued throughout the country. It is a curriculum- and standards-based educational and career planning tool that assesses students' academic readiness for college—and it's the capstone of ACT's College and Career Readiness System. Selecting from one of ACT's four In-School testing dates allows test forms to be available for the 2014-2015 testing cycle.

Custom Assessments

As previously mentioned, if the ADE chooses to implement a custom assessment for the SBCA, Pearson offers several distinct advantages that could help Arizona meet its 2014–15 operational goal, including an established Arizona assessment team, familiarity with Arizona's established assessment practices, in-house copies of AIMS operational and PARCC field test items, existing Arizona-specific systems that can be expanded to incorporate online test delivery, and pre-configured scoring capabilities for any AIMS and PARCC items used. Implementation of any interim, benchmark, or diagnostic assessments would likely not be possible until 2015–16 due to anticipated procurement timeframes and necessary lead times.

Requirement

c. Describe how, and the degree to which, the assessments are specifically aligned to Arizona's academic standards in mathematics and English language arts (reading and writing), include any alignment studies, if available.

Response

ACT Products

Alignment to Standards

The Common Core State Standards (CCSS), and, by extension, Arizona's College and Career Ready Standards, are a remarkable achievement, based in part on ACT's framework of college and career readiness, and developed with leadership from ACT content experts. The new ACT Aspire program capped by The ACT will support teachers and students with valid, actionable information to let them know where they stand. ACT envisioned not only a set of high-quality, vertically scaled summative assessments, but also a strong system that includes interim and classroom formative components to support teachers all year long. Results presented in the common language of Arizona's standards will help keep teachers aligned with their spirit and intent, while providing important state accountability measures.

ACT uses an extensive program of research for assessment development that will result in an empirically based, externally validated vertical scale measuring the construct of college and career readiness in each ACT Aspire content area. ACT's empirical understanding of the constructs allows for domain sampling in very specific ways to support accurate and reliable summative inferences of college and career readiness. Domain sampling occurs up and down the construct to acquire adequate information to predict success. ACT Aspire assessments have been crafted to carefully sample the domains effectively within a limited amount of testing time. This verifies that students do not lose valuable instructional days when they take summative assessments.

Traditional alignment methodologies have been based on content alignment. This can be item-to-standard alignment or standard-to-standard alignment. Current definitions of alignment have expanded to include comparing content coverage between an assessment and other curriculum documents. Alignment in the context of No Child Left Behind (NCLB) focuses on content agreement between state curriculum standards and state assessments. There are several well-documented methodologies in support of this definition, including the methodology developed by Norman Webb.

While summative assessments are critical for state accountability, no single summative test can address all standards in Arizona's College and Career Ready Standards. With this understanding, ACT is researching a new alignment methodology to address the complex challenge of aligning all ACT products to the CCSS and its derivatives.

Domain sampling, as a key component underlying ACT test blueprints, will drive alignment efforts to provide detailed information about ACT products and Arizona's College and Career Ready Standards. ACT recognizes that educational institutions must meet accountability requirements, and that such

requirements will often dictate an independent analysis as a prerequisite for commitment to a particular assessment.

The ACT is designed to cover the Common Core State Standards and Arizona College and Career Ready Standards. An internal alignment of the Arizona Academic Standards with ACT Course Standards conducted in 2010 indicated strong coverage in the English, math, science, and US history domains. ACT welcomes independent evaluations of its assessments as opportunities to discuss the goals and contents of its products.

Custom Assessments

Alignment

Because AIMS and national consortium items were written specifically to Arizona's College and Career Ready Standards, no additional cost or effort would be required to use them for Arizona's SBCA, other than perhaps an additional bias and sensitivity review for Arizona-specific cultural and ethnic issues. Other licensed third-party items would also require a content review to ensure alignment, preferably by committees of Arizona educators, per the ADE's established processes.

Requirement

d. Describe how the assessments' results can be compared to other states' criterion-referenced assessments expected to be in use beginning in the 2014-2015 academic year.

Response

ACT Aspire, ACT QualityCore, and The ACT

ACT assessments have an empirical link to college performance, and therefore the percentages of students meeting the college readiness benchmark on ACT assessments can be compared to the percentages of students meeting college readiness on other assessments.

ACT has been engaged in discussion with other criterion-referenced assessment providers, including the consortia, around creating linkages among assessments.

Custom Assessments

Comparability to Other States

As previously mentioned, by using items licensed from either of the two national assessment consortia, it would be likely that Arizona SBCA scores could be linked and compared to scores from other states using the full form of those consortia's assessments.

Requirement

e. Describe how the assessments are aligned to college / career expectations. Describe the validation process, including the role of post-secondary education in establishing the readiness expectations. Include any alignment studies, if available.

Response

ACT Aspire: Grades 3 through 10

College Readiness Standards and Benchmarks

Indicators of college and career readiness, proficiency, and other achievement levels are of utmost importance to Arizona's assessment system.

ACT Aspire encourages the approach whereby Arizona adopts ACT's College Readiness Benchmarks as indicators of college and career readiness and proficiency. This allows us to follow a student's academic process in the context of benchmark proficiency (four levels of readiness).

ACT's College Readiness Benchmarks provide a basis for these achievement levels, including empirical considerations such as one-year growth norms, alignment to NAEP and other current achievement levels in use, ACT's College Readiness Standards, and relationships of score points to probability of college success. The benefits to this approach include the following:

- **Reliance on ACT's Research.** ACT has documented the relationship of test scores to success in first-year credit-bearing college courses in English composition, college algebra, social science courses, and biology. The ACT College Readiness Benchmarks were updated this spring and are based on data from 214 colleges across the country, with more than 230,000 students. Recent ACT research also relates ACT scores to the probability of college degree attainment.
- **Continuity Across Grade Levels.** ACT's grade-specific readiness benchmarks link together empirically, to provide achievement levels harmonious across grade levels. For each ACT Aspire assessment, the readiness benchmarks will indicate whether students are on target to meet the ACT College Readiness Benchmarks.

ACT has capitalized on research partnerships with more than 100 postsecondary institutions for the ACT College Readiness Benchmarks (in English, mathematics, reading, and science), which indicate the scores needed to have at least a 50 percent chance of earning a "B" or higher grade in aligned first year credit-bearing college courses.

ACT College Readiness Benchmarks for earlier grade levels are defined as the scores needed to be on target to meet the ACT benchmarks by spring of grade 11. These benchmarks are monitored and updated as needed. ACT College Readiness Benchmarks for ACT Aspire are available for the 2013–14 academic year.

ACT Aspire will include a two-year "predicted path" that will indicate student performance based on average growth within the summative assessment. This growth prediction will be determined by growth norms for students assessed with ACT Aspire and The ACT.

Note that "one year of growth" determined by growth norms for students assessed with this model will allow for tracking of students progression of academic achievement against the ACT Readiness Benchmarks across grade levels.

In addition, this will allow schools and districts to easily follow a student's individual progress toward eventual attainment of the college readiness benchmarks provided in The ACT for grades 11 and 12.

For each assessment, achievement levels will be supported with descriptors of the knowledge, skills, and abilities possessed by students within each level. The achievement level descriptors will be derived empirically based on the performance of students on ACT Aspire and The ACT.

ACT QualityCore: Grades 9 through 11

On Course for Success: Defining a Syllabus-Driven Instructional System

During the 2003–2004 academic year, ACT and The Education Trust collaborated on a study to determine the courses, the level of rigor, and the instructional practices that are most likely to lead to success for students. Success is defined as meeting ACT's College Readiness Benchmarks for English, mathematics, or science that identify the level of knowledge and skills students need to attain in order to have a high probability of earning a grade C or higher in first-year college courses in English composition, algebra, or biology. The study focused on 10 schools, selected empirically, that are producing graduates who are meeting or exceeding the ACT College Readiness Benchmarks in proportions greater than those seen nationally. In nine of these 10 schools, the student population is made up of at least 40 percent minority-group membership and/or at least 50 percent of students with low-income status. The study, titled *On Course for Success* (ACT & The Education Trust, 2004), attempted to answer the question: What components of high school courses prepare students for successful entry into postsecondary education without the need for remediation?

Courses and teachers at the 10 schools were also selected by means of an empirical approach. Each of the participating schools was sent a list of its students who were ready for college—defined as meeting or exceeding one or more of the ACT College Readiness Benchmarks in two consecutive years. The schools identified the courses each student took and the teachers who taught those courses. The study team eventually focused on 69 courses and teachers with the highest percentages of students meeting the benchmarks. The team surveyed teachers about their educational experience, teaching philosophy, and instructional practices, and examined three weeks' worth of their lesson plans and instructional materials. Teachers were also interviewed and observed in their classrooms.

One of the four academic resources that the study found to be a crucial component of all the courses studied was high-level, college-oriented content. Successful students in these high schools were enrolled in college-preparatory courses and were learning the skills they need to be ready for college-level work. The content of these courses put students on a trajectory toward college from grades 9 through 12.

The study team's analysis of course syllabi, course descriptions, course content (including topics and/or specific standards covered), pacing charts, and instructional materials for these courses served as the foundation for the development of model syllabi for the following key high school courses in English, mathematics, and science:

- English 10, 11
- Geometry, algebra II, precalculus
- Biology I, chemistry, physics

The model syllabi represented syntheses of the materials, e.g., course content, course syllabi, pacing charts, and instructional materials the 69 participating teachers submitted. The course objectives derived from these materials were the first step in providing real answers to the question of what course rigor looks like and how it can best be taught. These objectives cover the important knowledge of a discipline, and the essential ideas that students must master if they are to have a grasp of the field. They also introduce knowledge and skills in an appropriate sequence, are of increasing intellectual sophistication at each higher level and/or course, represent what is manageable for instruction, and are measurable. These features enable the objectives to:

- Function as a set of organizing statements for an academic discipline
- Reflect a coherent body of scholarship
- Facilitate the development of lesson plans and units

Because this research is the only research of its kind in the US, the results of the research and the course standards played a major role in the development of the Arizona College and Career Ready Standards. The course syllabi, course descriptions, and course standards were then used to develop the remaining essential components of the core courses, including integrated curriculum and instruction, aligned formative and summative assessments, and supportive professional development for teachers and leaders.

Validation of Results: High School Survey

As a next step, we went beyond 69 classrooms to collect validity data from a wider sample of high-performing high schools nationwide. Lists of the course objectives were sent to a nationwide sample of English, mathematics, and science teachers chosen from a list of the 300 highest-performing high schools (i.e., schools in which significant numbers of students met one or more of the ACT College Readiness Benchmarks in two consecutive years). The teachers were surveyed about the importance and relevance of the objectives to college readiness. Teachers in each subject area were asked to indicate whether the knowledge and skills reflected in each objective are essential (that is, must be taught and mastered to prepare students for entry-level college coursework) and/or optional (may be taught to enrich a course), or unimportant and/or not applicable. The teachers were informed that the information they supplied would be used to identify the rigorous knowledge and skills taught in college-preparatory courses.

From the results of the survey, ACT staff refined the course objectives to reflect those that a majority of the responding teachers deemed to be essential. After reviewing all the survey responses, it was

determined that the most viable approach to this task was to include in these lists those objectives that a clear majority of the responding teachers in each content area agreed upon as being essential.

It is these final lists of course objectives that form the foundation for the test specifications of ACT's QualityCore end-of-course (EOC) examination program. ACT will continue to conduct research as the program evolves by gathering empirical evidence to confirm that ACT QualityCore test specifications remain at the leading edge of what constitutes a rigorous, integrated, instructional learning system in high-performing US high schools.

For the 12 ACT QualityCore courses, ACT provides an EOC Test Blueprint. These blueprints show how the test items are distributed across reporting categories and depth-of-knowledge levels (DOK). Each of the reporting categories within the EOCs is a logical, curriculum-based subgroup of the course standards for the course in question. Sample standards for each reporting category are provided, as well as definitions of the DOK thinking processes covered by the assessments. The constructed-response scoring guide and rubric are included. These resources support students' and teachers' understanding of the links between the course standards, the syllabus, and the exam reporting categories, making them as transparent as possible.

In order to support teachers' understanding of the examination system, ACT QualityCore Professional Development includes discussion of and practice with DOK levels, multiple examples of constructed-response tasks along with their scoring criteria and rubrics, and sample student responses. (Note: ACT QualityCore Professional Development is not included in the standard package's pricing).

Students' and teachers' understanding of the curriculum and the examinations is further supported by the tasks in the Formative Item Pools, which are aligned and coded to the ACT Course Standards and can be used to develop quizzes and other assessment-for-learning tools. There are multiple-choice and constructed-response items per course available for teachers to use at their discretion.

The ACT: Grade 11

One fundamental idea underlies the development and use of ACT's assessments: The best way to determine how well prepared students are for further education and for work is to measure as directly as possible the knowledge and skills needed in those settings and that can be supported by evidence. Accordingly, curriculum study in the content areas (English, mathematics, reading, and science) is ongoing at ACT and includes reviews of tests, curriculum guides, and national standards; a one-of-a-kind nationwide survey of educational practices and expectations; and meetings with content experts. Information and evidence from these sources defines the scope and sequence for each of the areas measured by ACT assessments.

The ACT National Curriculum Survey®, conducted by ACT every three to four years, continues to play a primary role in guiding the development of ACT's assessments. ACT surveys thousands of elementary school, middle school/junior high school, secondary, and postsecondary teachers in English/writing, reading (including English language arts and social studies teachers), mathematics, and science for the purpose of determining what skills and knowledge are currently being taught and which are considered important for success at each grade level for college and career readiness. The survey data ACT collects about what entering college students need to know and be able to do to be ready for college-level

coursework helps inform ACT's assessments and confirm that they meet the needs of college and career readiness.

Requirement

f. Describe the available accessibility features, as well as assessment accommodations for individuals with disabilities and English Language Learners.

Response

ACT Aspire

Accommodations and Accessibility

Making common statewide assessments accessible to virtually all students is a shared concern for Arizona, Pearson, ACT Aspire, and ACT. ACT's attention to accessibility is one factor that makes its ACT Aspire assessments especially suitable to serve Arizona's needs.

ACT Aspire will provide Arizona students with accessible content on the ACT Aspire website as well as in its testing programs. This content, along with the ACT Aspire online delivery platform, will adhere to specific guidelines and legislation, as described below, with usability as the primary standard by which it judges the accessibility of its web content and test materials.

US Rehabilitation Action Section 508

All online system content for ACT Aspire will meet Section 508 accessibility guidelines. ACT expects to be compliant with the expected revisions to Section 508 Standards within one year of their release.

Accessible Portable Item Protocol and Personal Needs Profile

Both ACT and Pearson are contributing members of the IMS Global Learning Consortium (GLC) and active members of the Accessible Portable Item Protocol™ (APIP) Working Group. IMS GLC members provide leadership in shaping and growing the learning industry through community development of interoperability and adoption practice standards. The Question and Test Interoperability™ (QTI)/APIP Alliance is a community of suppliers, vendors, and organizations committed to the development and adoption of interoperability standards. The Alliance offers an online engagement with experts from other member organizations, IMS staff, and a library of resources shared by the participants.

In addition, Wayne Ostler, Pearson Vice President of Technology Strategy, and currently co-chair of the APIP Working Group, has used his position as Assessment Lead on the Schools Interoperability Framework (SIF) Association Assessment Working Group to help spearhead efforts to align the APIP and SIF specifications to eliminate overlap and stimulate collaboration between the two standards.

ACT involvement in IMS work and in the APIP standards also continues to increase, with ACT staff representation in the APIP Working Group and the APIP End Users group, among others. ACT and Pearson are working jointly to incorporate APIP and Personal Needs Profile (PNP) standards into test

authoring and online delivery systems in conjunction with APIP support in the delivery platforms over the next year.

Braille and Large-Print Administration of ACT Aspire

To meet Arizona's student testing needs, Pearson and ACT will produce sufficient braille and large-print materials for each test administration window. We also will provide appropriate manipulatives to accompany braille and large-print testing materials. Working as we have in the past with the ADE's braille reviewers, Arizona would have final approval of the braille transcription service and braille materials used.

Arizona will have available the following paper-based formats for all assessments:

- Regular print paper and pencil
- Large type paper and pencil
- Braille and tactile graphics

Students may also respond verbally to a scribe and may have options for extended time and multiple breaks. [redacted]

Braille forms, integrated with the required tactile graphics for each test form, currently provide a smooth testing experience for braille users. The standard procedure for ACT Aspire is to have the student dictate the answer to a scribe, who transcribes the response onto a scannable answer document.

Assistive Technology

Arizona students will benefit from ACT Aspire assessments that support a range of accessibility tools and formats, including but not limited to, the following:

- Audio supports, such as text-to-speech and verbal descriptions of graphics
- Visual supports, such as various color contrast settings, including reverse contrast, pastel background choices, color overlay, screen magnification, line reader, and highlighter
- Closed captioning of video content
- Speech-to-text for some components

Online Accommodations			
Accommodation	**Description**	**Notes**	**Valid Options**
Auditory Calming	Selected sounds play as the student moves through the test.	Any number of sound files	Any number of sound files
Zoom	Vision-impaired students can magnify the entire screen up to 300 percent		Enabled/Disabled
Color Contrast	Students can changes the appearance of test content by changing the text and screen colors or contrasts.	▪ Accessibility panel allows for contrast options including filters, background/text combinations, and overlays ▪ Customizable background/text combinations	Enabled/Disabled
Text-to-Speech (TTS)	A spoken voice recites test content. The text is visually tracked on screen as it is being recited.	▪ Word by word highlighting ▪ Can change the rate of speech (once set, settings stay on from item to item) ▪ Can be set to let student set speed or it can be pre-determined. ▪ Highlight colors can be specified	▪ Enabled ▪ Disabled ▪ Audio Speed ▪ Highlight colors ▪ Word-based Click to Hear ▪ Can be set for students who do not have a visual impairment (read-along version); assumes students can see and interpret graphic or non-text content ▪ Can be set for students who are blind or have a visual impairment; includes keyboard navigation capabilities and text descriptions of graphics and non-text content
Answer Masking	Students can block out all answer choices.	▪ Student can remove mask for each answer choice ▪ All masks can be reset by student	▪ Enabled ▪ Disabled ▪ Color
Line Highlighter/ Reader	Students can move a highlighter up and down to focus on a line.	Draggable	▪ Enabled ▪ Disabled ▪ Color and Opacity ▪ Height
Sign Language	American Sign Language translations are available.	Pre-recorded human interpreters or programmable avatars (TBD)	▪ Instructions only ▪ Instructions and test content

Providing a Range of Accommodations. The ACT Aspire package includes provisions to accommodate a range of student needs.

ACT's standard test administration instructions state that for students with specified accommodations who are authorized to take the test with extended time, the student must normally complete each subject session individually and may not return to a previous subject test after completion. However, as required by a student's IEP, this can be adjusted as required.

Within the student data enrollment section of the online platform, schools will be able to enter information related to each student's Personal Needs Profile (PNP), which will enable the system to deliver the required formats or accessibility features for that student. That PNP system will not only meet APIP standards, but can enable accommodation coding, if Arizona desires.

ACT QualityCore

Accommodations

In addition to ACT QualityCore's standard paper-and-pencil and computer-based testing formats, ACT QualityCore tests and forms are available in large-print, braille, reader's script and audio CD formats to address the accommodation needs of students with disabilities.

Special Testing—Administration Types

ACT QualityCore provides materials for standard and special administrations. Students with physical or learning disabilities who cannot complete the ACT QualityCore tests in the standard time limits using the standard materials may be tested, at the school's discretion, under special conditions and/or using special testing materials available from ACT. Special test materials include braille and large print (18-point) test books, large print answer sheets, audio CDs, and reader's scripts.

It should be noted that:

- Some accommodated materials, e.g. braille, may differ from the standard administration test form.
- The online reports will not display the accommodations code gridded on the answer sheet.

Accommodations Process and Materials				
Topic	**Braille**	**18-Point Large Print**	**Audio CD**	**Reader's Script**
Ordering	▪ Client contacts ACT Customer Services for the Special Testing Materials order form to be faxed in. Order must be placed no later than four weeks prior to test	▪ Client contacts ACT Customer Services for the Special Testing Materials order form to be faxed in. Order must be placed no later than four weeks prior to test	▪ Client contacts ACT Customer Services for the Special Testing Materials order form to be faxed in. Order must be placed no later than four weeks prior to test	▪ Client contacts ACT Customer Services for the Special Testing Materials order form to be faxed in. Order must be placed no later than four

Accommodations Process and Materials				
Topic	**Braille**	**18-Point Large Print**	**Audio CD**	**Reader's Script**
	administration start date.	administration start date	administration start date	weeks prior to test administration start date
Shipping	▪ Shipped in separate boxes to arrive with other test materials. ▪ District/school retains boxes, reusing boxes to return the Braille materials.	▪ Shipped in separate boxes to arrive with other test materials. ▪ District/school retains boxes, reusing boxes to return the large print materials.	▪ Shipped in separate boxes to arrive with other test materials. ▪ District/school retains boxes, reusing boxes to return the large print materials.	▪ Shipped in separate boxes to arrive with other test materials. ▪ District/school retains boxes, reusing boxes to return the large print materials.
Security bar codes	▪ One unique bar code for each booklet, with the understanding that one assessment will have multiple booklets	▪ One for each booklet	▪ One for each CD	▪ One for each reader's script

Accommodations Process and Materials				
Topic	**Braille**	**18-Point Large Print**	**Audio CD**	**Reader's Script**
Student answer document	Test Administrator: ▪ Transcribes student answers to student answer document. ▪ Indicates the accommodation on the student answer document. ▪ Submits transcribed, scannable answer documents with other completed, standard administration answer documents.	▪ Students record their answers directly in the large-print answer document. ▪ Test Administrator: Writes the student's name, teacher, school, and district on the front cover of the large-print ACT QualityCore student test booklet. ▪ Transcribes answers to standard administration answer document. ▪ Indicates the accommodation on the student answer document. ▪ Submits transcribed, scannable answer documents with other completed answer documents from the student's class(es).	Test Administrator: ▪ Transcribes student answers to scannable student answer document. ▪ Indicates the accommodation on the student answer document. ▪ Submits transcribed, scannable answer documents with other completed answer documents from the student's class(es).	N/A

Accommodations Processes. ACT Aspire uses efficient processes for implementing accommodations.

Since its inception in 1959, ACT has designed its exams for ease of use by test takers. ACT believes that the original designers of The ACT, by establishing a good test design, observed in advance the seven principles of Universal Design as identified by the Center for Universal Design.

In June 2002, when the National Center on Educational Outcomes, in collaboration with the Council of Chief State School Officers and the National Association of State Directors of Special Education, published its Synthesis Report 44 on Universal Design Applied to Large Scale Assessments (Thompson, Johnstone, & Thurlow, 2002), ACT reviewed its exams' design and procedures in light of that report. We found that the original designers of the ACT had succeeded in designing for a diverse audience from the

beginning, and that consequently the seven principles of Universal Design have been integral to daily editorial decisions in ACT's test development of ACT QualityCore.

That report, building on seven principles, identifies "seven elements of universally designed assessments." These, with ACT's commentary regarding ACT QualityCore assessments, are as follows:

1. **Inclusive Assessment Population.** ACT considers students in developing the ACT QualityCore assessments. Assessments undergo rigorous reviews both in-house and from independent consultants before being added to the formative assessment pool or placed on an operational EOC test.
2. **Precisely Defined Constructs.** The ACT QualityCore assessments are based upon empirically-derived course syllabi and course standards. The ACT QualityCore assessments are strictly tied to the course syllabus and course standards.
3. **Accessible, Non-Biased Items.** ACT works with teachers to develop the assessment items, with sensitivity to a diverse examinee population. ACT QualityCore assessment items are submitted to independent reviewers for content and for fairness. The item-response data from operational administrations are analyzed for differential item functioning (DIF) through the use of the standard difference in proportion correct (STD) and the Mantel-Haenszel common odds-ratio (MH). Fewer items have been flagged by these DIF analyses than would have been expected by chance alone. The items that are flagged are analyzed to determine if any source of actual bias went undetected by the item review process. To date, no source of actual bias has been found.
4. **Amenable to Accommodations.** Because the ACT QualityCore assessments have been designed from the outset according to high publishing-industry standards, in keeping with curricular materials, and in consultation with in-service educators, the tests have been amenable to accommodations across the range of testing populations, conditions, and formats. We keep tests as simple and straightforward as possible, consistent with curricular requirements—and this applies equally to vocabulary, graphics, typographic design, page layout, and the interrelationships among all the elements.
5. **Simple, Clear, and Intuitive Instructions and Procedures.** ACT QualityCore assessments comply with the Standards for Educational and Psychological Testing (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999, p. 47) with respect to test instructions that are simple, clear, consistent, and understandable, as well as guidelines and procedures for test administration.
6. **Maximum Readability and Comprehensibility.** Rather than relying on readability formulas (whose mechanical application has been shown not always to be in students' best interest), ACT relies instead on teacher involvement, expert opinion, and consultants' commentary to gauge the grade-level and content-area appropriateness of the ACT QualityCore assessments. The 10-point "Test Item Readability checklist" given on page 15 of "Universal Design Applied to Large Scale Assessments," despite minor verbal differences, is essentially a subset of the longer list of guidelines that ACT employs in its quality assurance procedures for the development and review of test materials:

1. **"Students would likely have the experiences and prior knowledge necessary to understand what the question calls for."** The research-based course syllabi and course standards of the ACT QualityCore program and ACT's test development procedures assure that the ACT QualityCore assessments are linked to school curricula and reflect the diversity of students' experiences.
2. **"The vocabulary is appropriate for the intended grade level."** The engagement of grade-level-appropriate teachers in the test development process, together with ACT test development guidelines and extensive consultant reviews, provides appropriate vocabulary.
3. **"Sentence complexity is appropriate for the intended grade level."** ACT's selection of item writers, test guidelines, and detailed reviewing steps provide for appropriate sentence complexity.
4. "**Definitions and examples are clear and understandable."** ACT so specifies in its Item Writer Guides and internal editorial standards, and internal and external test reviewers monitor this aspect of our tests.
5. **"The required reasoning skills are appropriate for the students' cognitive level."** The research-based course syllabi and course standards of the ACT QualityCore program, together with ACT's test development procedures, see that the ACT QualityCore assessments require reasoning skills appropriate for the students' cognitive level.
6. **"Relationships are made clear through precise, logical connectives."** ACT editorial standards together with the expertise of the teachers who write the tests and the ACT staff who edit them, confirm that the test text is logical and that relationships are made clear. Close reviews by content experts confirm that ACT meets high standards of composition and that test text is in line with current standards and usage in each area tested.
7. **"Content within items is clearly organized."** ACT editorial standards, together with the expertise of the teachers who write the tests and the ACT staff who edit them, verify that the content within items is clearly organized. Close reviews by content experts confirms that ACT test items meet high standards for clarity and logic and that the items conform to high professional standards for clarity and usage in each area tested.
8. **"Graphs, illustrations, and other graphic aids facilitate comprehension."** ACT provides graphic aids for the ACT QualityCore assessment stimuli and items whenever such aids are appropriate. Some of these graphic elements are adapted from published sources, others are developed especially for the tests. Graphics are tailored for the stimulus or item they illustrate and for the grade level of the examinees, and are subject to the same close review by content experts as is the text.
9. **"The questions are clearly framed."** The expertise of ACT's item writers and staff and ACT's experience over decades of test development combine to frame the ACT QualityCore assessment questions clearly. ACT's staff is assisted by in-house guides to item development and by the comments and suggestions of the many experts who participate in ACT's content and fairness reviews.

10. **"The content of items is of interest to the intended audience."** ACT is sensitive to the requirement that test content be interesting to students and takes care to select materials that are interesting to examinees. ACT emphasizes to item writers that test content be interesting; periodically surveys examinees to find out what test materials were the most and least successful in gaining their interest; and attends to the advice of teachers, fairness consultants, and content experts as to what topics and approaches are, and are not, appropriate for each test at each grade level.

    In sum, ACT makes every effort to maintain "maximum readability and comprehensibility" in our ACT QualityCore assessments. Throughout test development, we aim to use plain language to focus every test item on the knowledge and skill being tested—that is, to avoid any material extraneous to what is being tested—in the interest of avoiding what has been termed "construct-irrelevant variance" in measurement.

7. **Maximum Legibility.** Universal Design Applied to Large Scale Assessments conveniently treats legibility factors under three major headings: Legible Text; Legible Graphs, Tables and Illustrations; and Legible Response Formats. ACT has carefully weighed the many requirements of its test formats and has achieved what we believe to be an optimal balance of features. While not every feature may be ideal, the design as a whole has the optimal features for the majority of test takers.

    ACT provides a comprehensive manual that outlines and describes how to administer the four available accommodations for ACT QualityCore:

    - Braille
    - Audio CD
    - Large Print
    - Reader Script
    - Extended-time is also available for ACT QualityCore examinees, whether testing by paper-and-pencil or the computer-based platform.

ACT QualityCore provides materials for standard and special administrations. Students with physical or learning disabilities who cannot complete the ACT QualityCore tests in the standard time limits using the standard materials may be tested, at the school's discretion, under special conditions and/or using special testing materials available from ACT. Special test materials include braille and large print (18-point) test books, large print answer sheets, audio CDs, and reader's scripts.

The ACT

To provide that official ACT scores reported to colleges and other entities from statewide testing are comparable to scores earned through other forms of ACT testing involving the application of ACT's test accommodations policies, ACT supports the following two forms of accommodations on the ACT when it is administered as part of state testing:

- **ACT-Approved Accommodations** result in ACT scores that are fully reportable to colleges, scholarships, and other entities. Only students with professionally diagnosed and documented disabilities who receive accommodations in school should apply for ACT-approved accommodations.
- **Optional State-Allowed Accommodations** result in ACT scores that are not college reportable; they are used for state accountability purposes only. Limited English Proficiency (LEP) students who do not have a disability but receive accommodations in school should request state-allowed accommodations.

Requesting Accommodations on The ACT

In general, all accommodations on The ACT must be requested and reviewed by ACT. However, there are limited exceptions. For example, because testing will normally occur at the local school rather than a separate test center, some arrangements do **not** require review or prior approval from ACT (e.g., placement at the front of the room). Such arrangements are noted on the accommodations summary table at the end of this response as "local decision" meaning they do **not** require ACT review or approval.

Participating schools must appoint a Test Accommodations Coordinator (TAC) who will submit requests for accommodations to ACT. The TAC has access to two different forms specifically designed for a state testing administration of the ACT:

- **ACT Request of ACT-Approved Test Accommodations.** This form is used to request ACT approval of accommodations for students who meet ACT eligibility requirements. See information about ACT's review of these requests in the next section below.
- **ACT Application for State-Allowed Accommodations.** This form is used to order test materials for students who will test with State-Allowed accommodations. These students are those who do not meet ACT's eligibility requirements (e.g., LEP students with no disabilities) or whose requests for ACT approval have been denied. ACT will ship the materials ordered for each student; no review or approval process will be conducted.

ACT Review of Requests for ACT-Approved Accommodations

ACT will review requests for ACT-Approved Accommodations by applying the Americans with Disabilities Act (ADA) standards used for all such requests. Attached is a draft of the ACT Accommodations Summary Table. ACT will work with the each state agency to explain and refine responses listed in the accommodations summary table to tailor to the needs of students testing with accommodations. Note that not every request for an accommodation listed on the attached accommodations summary table as available will be approved for a particular student. Approval is dependent on submission of all required documentation by the stipulated deadline and review by ACT. It is possible for ACT to approve an accommodation for one student, while the same accommodation may be denied for a different student. ACT's decision whether to approve the requested accommodations under the ADA will determine whether resulting ACT scores can be reported to colleges **in addition** to being used for state accountability purposes. Note also that accommodations approved for multiple choice tests are approved for use on all tests in the battery.

Ordering State-Allowed Accommodations Materials for The ACT

Students who do not meet ACT eligibility requirements (e.g., LEP students with no disabilities) or whose requested accommodations are denied by ACT have two options: 1) Test under standard conditions and receive college reportable ACT scores, or 2) submit an order for "State-Allowed" accommodations materials resulting in ACT scores that are NOT college reportable.

Important Note: TACs must submit an order for "State- Allowed" accommodations for each applicable student so that ACT can ship the correct ACT test materials—which are **different** from those used by students testing with ACT-approved accommodations.

ACT scores resulting from testing with "state-allowed" accommodations are **not** college reportable, but may be used for state accountability purposes. Thus, some students will achieve ACT scores that are college reportable because their accommodations have been approved by ACT, while others using the same accommodations will achieve ACT scores that are not college reportable because their use of those accommodations was not approved by ACT.

[redacted]

The following figure summarizes accommodations for The ACT.

Summary of Accommodations for the ACT				
Accommodation	**ACT Equivalent**	**May Request**	**ACT Comments**	**College Reportable Scores**
Timing/Scheduling				
Time of day for testing altered	Local decision	Yes	The ACT for state testing purposed is conducted in a 2-week testing window for students testing with accommodations. Day of the week or time of day is determined locally.	Yes
Additional breaks provided	"Stop-the-clock" breaks	Yes	Applies to students who are diabetic or have other medical conditions requiring frequent breaks. Normally available only with standard time. If requested with extended time, must provide documentation to support need for "stop-the-clock" breaks in addition to extended time.	Yes
Test administered in several sessions	Testing over multiple days	Yes	For certain formats and disabilities, ACT will assign a timing code for the ACT based on the test format and disability, up to triple time. Testing over multiple days may be approved, but each test must be completed during a single session.	Yes
Longer breaks between test sessions	Extended time	Yes	Extended time testing is normally time-and-one-half in a single session using regular or large print.	Yes
Longer breaks between test sessions	Testing over multiple days	Yes	For certain formats and disabilities, ACT will assign a timing code for the ACT based on the test format and disability, up to triple time. Testing over multiple days may be approved, but each test must be completed during a single session.	Yes
Extended testing time for test sessions	Extended time	Yes	Extended time testing is normally time-and-one-half in a single session using regular or large print. For certain formats and disabilities, ACT will assign a timing code for the ACT based on the test format and disability, up to triple time. Testing over multiple days may be approved, but each test must be completed during a single session.	Yes

Summary of Accommodations for the ACT				
Accommodation	**ACT Equivalent**	**May Request**	**ACT Comments**	**College Reportable Scores**
Response				
Lined paper turned sideways to help align math problems	Visual aids provided by school or student	Yes	Provided by school or student; student must test individually.	Yes
Use of a low-tech assistive writing instrument	Device provided by school or student	Yes—submit details with request	Assistive technology must be approved by ACT, and must not compromise test security.	Yes
Alternative indication of response (i.e., circle, point to, state, or otherwise indicate answer choice)	Local decision (mark response in test booklet)	Yes	Test Accommodations Coordinator must arrange to transfer the response to the answer folder in the student's presence after testing.	Yes
Alternative indication of response (i.e., circle, point to, state, or otherwise indicate answer choice)	Use of scribe	Yes	Testing staff is responsible for marking the student's answer choices on the answer document and/or transcribing the essay. The student must indicate punctuation, format, and spell all key words for Writing. For college reportable ACT scores, audio documentation of test session must be returned to ACT.	Yes—only if recording of test session returned to ACT
Use of an approved, bilingual word-to- word dictionary	Use of bilingual word-for-word translation glossary	Yes—State-Allowed only	Provided by school or student for LEP students.	No
Use of a computer or other assistive technology (AT) device	Computer or AlphaSmart for writing test only	Yes	Provided by school or student. Applies only to ACT Writing Test. ACT instructions for printing and returning essay must be followed precisely.	Yes

Summary of Accommodations for the ACT				
Accommodation	**ACT Equivalent**	**May Request**	**ACT Comments**	**College Reportable Scores**
Setting and Environment				
Preferential seating	Local decision	Yes	Preferential seating is determined locally.	Yes
Special lighting conditions	Local decision unless requesting off-site or required by approved accommo-dation	Yes	If setting is off-site, appropriate off-site application must be approved by ACT. Note that individual testing is required for selected accommodations (e.g. if approved accommodations could disturb others or if approved for a reader).	Yes
Small group setting	Local decision unless requesting off-site or required by approved accommo-dation	Yes	If setting is off-site, appropriate off-site application must be approved by ACT. Note that individual testing is required for selected accommodations (e.g. if approved accommodations could disturb others or if approved for a reader).	Yes
Test individually	Local decision unless requesting off-site or required by approved accommo-dation	Yes	If setting is off-site, appropriate off-site application must be approved by ACT. Note that individual testing is required for selected accommodations (e.g. if approved accommodations could disturb others or if approved for a reader).	Yes

Summary of Accommodations for the ACT				
Accommodation	**ACT Equivalent**	**May Request**	**ACT Comments**	**College Reportable Scores**
Presentation				
Allow use of special furniture or equipment for viewing test	Varies	Yes—submit details with request	Requests are considered individually based on documentation submitted.	Yes—depending on the details submitted with the request
Student reads aloud to him or her self	Local decision	Yes	Student must test individually.	Yes
Use of assistive technology to magnify/enlarge	Visual aids provided by school or student	Yes—submit details with request	Device may range from magnifying glass to CCTV. Depending on device may require individualized testing.	Yes
Use of assistive technology to magnify/enlarge	Large-type test booklet	Yes—submit details with request	ACT produces large type version 18 pt. font. Large type may be used with standard or extended time. It is reserved for students with visual impairments.	Yes
Use of acetate film	Local decision	Yes	Provided by school or student.	Yes
Access to sound amplification system	Local decision	Yes	Provided by school or student; student must test individually if device used causes distraction.	Yes
Access to a large print version of test	Large type	Yes	ACT produces large type version 18 pt. font. Large type may be used with standard or extended time. It is reserved for students with visual impairments.	Yes
Access to own resources (i.e., bold print protractor, real coins, bold/raised line graph paper, bold/raised line writing paper)	Varies	Yes—submit details with request	Requests are considered individually based on documentation submitted.	Yes—depending on the details submitted with the request
Braille test format	Braille	Yes	ACT produces Braille. The Braille format may require extended time of up to triple time.	Yes
Access to an interpreter for sign language	American Sign Language (ASL) for directions only	Yes	American Sign Language (ASL) may be used only for spoken instructions exactly as provided in the administration manual. ASL for test items is not ACT-Approved.	Yes—ASL Directions Only
Access to an interpreter for sign language	Exact English Signing (EES)	Yes	Exact English Signing (EES) of test items may be requested and approved in specific cases for college reportable scores.	Yes—EES

Summary of Accommodations for the ACT				
Accommodation	**ACT Equivalent**	**May Request**	**ACT Comments**	**College Reportable Scores**
Access to a talking/screen reading device (can NOT be used for reading comprehension portion of test)	Audio DVD	Yes	ACT Oral Presentations include all test questions and answers as well as any passage associated with those items. ACT-produced audio version (Audio DVD) must be used or ACT-produced reader's script read verbatim in English. For college reportable ACT scores, students must test individually if not using audio version with headset.	Yes
Test read aloud by test administrator (except items testing comprehension)	Reader Script	Yes	ACT Oral Presentations include all test questions and answers as well as any passage associated with those items. ACT-produced audio version (Audio DVD) must be used or ACT-produced reader's script read verbatim in English. For college reportable ACT scores, students must test individually if not using audio version with headset.	Yes
Reading all assessment directions in student's native language ▪ Student must be dominant in that native language; **and** ▪ Student's English proficiency is determined to be basic or lower intermediate; **and** ▪ Student receives bilingual instruction in their native language for the maintenance of that language	Yes	Yes (State-Allowed only)	Includes spoken instructions and directions printed in the test booklets. If student's reason for accommodations is limited English proficiency, student must order "State-Allowed" accommodations materials.	No

Summary of Accommodations for the ACT. In general, all accommodations on the ACT must be requested and reviewed by ACT. ACT will work with the ADE to explain and refine responses listed in the accommodations summary above to tailor to the needs of students testing with accommodations.

Note: Unless an accommodation is specified as applying to the writing test only, the accommodation applies to all tests.

Requirement

g. Describe any practice and/or sample assessments that are available.

Response

ACT Aspire Exemplars

The ACT Aspire solution includes exemplar tests that students and educators can use to become familiar with the type of items used on the ACT Aspire assessments and the navigation features of the Pearson TestNav system. The exemplar experience is intended to replicate the live TestNav experience and environment. As a result, once responses are submitted, a user cannot go back into the test to change responses. Users can launch the exemplars as often as they prefer. The exemplars can be located at http://www.discoveractaspire.org/pages/test-features.The following figure outlines the available exemplars by subject across grades 3–10:

Subject	Number
Math	30
Science	43
Reading	36
English	24
Writing	7
Total	140

ACT QualityCore

Under the ACT QualityCore program, sample EOC tests for English 10, algebra I, biology, and US history are available in PDF format.

The ACT

Test Preparation Resources

The chart below outlines the practice test and other test preparation resources ACT makes available to students, parents, and educators. Many of these resources are offered at no charge.

ACT's Test Preparation Resources

Title	Media	Distribution	Other
Preparing for the ACT	Printed booklet Complimentary PDF download available at www.actstudent.org	ACT distributes quantities of booklets to ALL high schools annually. A PDF is available on the ACT student and ACT educator websites. An order form for additional hardcopies is available on the educator website.	The document includes tips for taking the test, information about types of calculators allowed on the ACT Mathematics Test, and a practice test with answer key.
Online Practice Test Questions	Available at www.actstudent.org	Online only, no charge. A practice DVD (for accommodation testing) is available. TACs will receive information about ordering the DVD in their establishment packet of materials.	Students have access to online practice questions. For each question, students click on an answer to see if they are correct. Explanations of both correct and incorrect answers are given. Five sets of questions are available for English, Mathematics, and Science; four sets are available for Reading.
ACT Online Prep™	Available for purchase at www.actstudent.org	For fee, internet-delivered, test preparation program designed by ACT test development.	Features a Diagnostic Test, two ACT Practice Tests, practice for the optional ACT Writing Test, and comprehensive review sections for each of the four required test tests, as well as the Writing Test.
The Real ACT Prep Guide – with CD	Printed book available for purchase at www.actstudent.org	For fee, test preparation book based on real ACT tests.	This guide provides access to three real ACT Tests, including samples and practice tips for the optional Writing Test. The CD contains supplemental resources for choosing a major, career planning, exploring colleges, and college planning advice.
Sample Test Booklets	Printed book available for purchase through ACT customer service or on an ACT-provided order form.	For fee. Available to schools with a valid ACT high school code.	Retired ACT forms may be purchased by schools. Included in each booklet is a sample answer folder and scoring keys. Two forms of the ACT are reserved for use by schools, while two forms are available to test preparation organizations.

Test Preparation. ACT offers a wealth of test preparation resources, many offered at no charge.

Custom Assessments

Practice Tests

Pearson's online testing system for custom assessments includes a separate training site to host practice tests containing representative samples of released items. The training site is typically activated well in advance of operational administrations, which provides users with a system to familiarize themselves with the online environment.

Through the training site, test administrators and technology coordinators can become familiar with the online testing experience before live testing. Teachers can practice adding students, updating students, setting up sessions, and monitoring students while taking the practice tests. Technology staff can practice configuring TestNav and using proctor caching before the live assessment.

Students can practice testing using the sample form(s) in the training site, as assigned by local test administrators, or through standalone electronic Practice Assessment Tests (ePATs), which offer the following features:

- A practice version of each assessment that simulates the live test experience
- Students can practice using online tools
- Students can practice interacting with and responding to items

ePATs are typically built using the practice test items, and students can take them any time from any Internet-connected computer, at home with their parents or at school. ePATs use the same interface, layout, and tools as the live test, but students can practice on their own time, as often as they want, with no pressure. Students with any level of computer skills can become familiar with TestNav before a high-stakes online test.

The use of ePATS as an important part of student test preparation. ePATS are easy to access and give the students an opportunity to the use the same interface they will be using during operational testing. Thus far in 2013, Pearson has delivered over 3.8 million ePATs for our online testing customers.

3. Computer-Based Assessments

Requirement

a. If applicable, describe the computer-based option(s) for the administration of the assessments.

Response

Delivering Arizona Assessments Online

Arizona educators will welcome the experience and confidence that Pearson brings to large-scale, high-stakes online test delivery. Since 2001, some of the nation's largest assessment programs have been using our test delivery engine, TestNav, to make the transition from paper-based to online testing.

System failure would damage the trust of Arizona stakeholders, so we design our scalable online testing platform to perform reliably with a variety of school infrastructures. Pearson has processes for verifying that Arizona schools and Pearson servers will be ready on test day. The proctor caching software that Pearson provides will enable Arizona districts with low-bandwidth Internet service to enjoy rich assessment content with high-bandwidth performance. In the rare event of a system interruption, TestNav saves an encrypted backup file of responses, so student effort and confidence are not lost.

By choosing a proven leader in online testing, Arizona can take advantage of Pearson's assessment experience and economy of scale. In the first eleven months of 2013, we delivered more than 11 million online tests for eight different state assessment programs, including the top online testing programs in the nation: Florida, Texas, and Virginia.

Similar evidence of our test delivery success in 2013 includes the following:

- 205,000 tests in a single hour
- 429,000 tests in a single day
- 1.2 million tests in a single week

Online Delivery Using TestNav

Computer-based test delivery presents tremendous opportunities to develop innovative assessment items that allow students to more fully demonstrate their abilities. Innovative items allow students to show how they arrived at an answer, and scoring is permitted along a range of possible point values based on how well they understand a particular process, rather than simply scoring their answers as only right or wrong.

Oftentimes, simple online items that students cannot guess, such as fill in the blank, are referred to as "innovative." While such item types do add somewhat to the depth of knowledge and ability measured, they fall far short of the potential that computer-based delivery offers, and they fail to live up to the expectations students and teachers using online instruction will bring to their online testing experience. Pearson offers innovative items with greater depth in terms of appearance, action, tools and manipulatives, and ability to demonstrate what a student knows and can do.

Pearson has been working with technology-enhanced (TE) items for a number of years. Our team members are familiar with the practical challenges associated with scoring, usable design, and bandwidth requirements that these item types can present. The complexities of technology-enhanced items require a vendor with experience and staff skilled in the areas of interactive content design, usability engineering, and software development. Pearson has developed highly collaborative groups that work within these areas to develop technology-enhanced items and content/technology solutions, such as simulations, and new interfaces and tools appropriate to particular subject areas.

Over the past year, we published a series of white papers and guides addressing key topics and issues relating to online testing, including Assessment Technology Standards, Next-Generation Assessments Interoperability Standards, Considerations for Next-Generation Assessments: A Roadmap to 2014, and Cognitive Lab Evaluation of Innovative Items. These papers are available at www.pearsonassessments.com/nextgeneration. Additional papers that we have published in the area of online testing are available at http://www.pearsonassessments.com/research under "Research Reports."

Technology enhanced items are a key part of our nation's shared vision for a next-generation assessments. Since 2005, Pearson has been a leader and innovator developing and delivering TE items. We have implemented, or are currently implementing, multiple programs using TE items. These projects include the following:

- Minnesota Comprehensive Assessments (MCA)—Science and Math
- Virginia Standards of Learning (VASOL)—Math, Science, and English Language Arts
- Virginia Modified Achievement Standards Test (VMAST)—Math and English Language Arts
- Virginia Algebra Readiness Diagnostic Test (ARDT)—Algebra
- Oklahoma Core Curriculum Tests—Math
- Oklahoma End of Instruction (EOI)—Algebra I, Geometry, Algebra II, US History, Biology I, English II, English III
- American Diploma Projects (ADP) Assessment Consortium—Algebra I, Algebra II
- Texas Assessment of Knowledge and Skills (TAKS)—Science
- Texas End of Course (EOC)—Chemistry, US History, Biology, Geometry
- Readypoint Nursing—scoring of TE items within a computer adaptive test
- ACT Aspire—Reading, English, Writing, Mathematics and Science
- Colorado Assessment Program—Science and Social Studies

- GED Testing Program—Literacy, Mathematics, Science, and Social Studies

As previously mentioned, Pearson and ACT create items using the IMS Global Question & Test Interoperability Specification (QTI), the same standard adopted by PARCC. The table below identifies QTI item types that Pearson is working on or that are on the roadmap and will be available before spring 2015. This list is not exhaustive as we are continuing to build out item type capabilities using the QTI specifications.

QTI Interaction Types Supported by TestNav	
Choice	▪ **Multiple Choice** Image, math, or text-based choices with a single choice made using a radio button. ▪ **Multiple Choice, Multiple Select** Image, math, or text-based choices with multiple selections possible using checkboxes. ▪ **Inline Choice, Multiple** Multiple drop-down menus are used within a range of text to provide several options consisting of words or phrases. Each drop down is its own interaction. ▪ **Multiple Choice (Cloze), Single** Test-taker may choose which word or phrase should be filled into a blank within a sentence or paragraph or image
Order	▪ **Ordering** Test-taker drags elements into position to indicate order. Draggers can be text or image and can be placed horizontally or vertically.
Text Entry	▪ **Textual Response, Simple (Fill in the Blank)** A box for typing text is provided to the test-taker; valid characters and the maximum amount of characters can be defined ▪ **Textual Response, Simple (Gridded Response)** A box for typing text is provided to the test-taker; valid characters and the maximum amount of characters can be defined
Extended Text	▪ **Textual Response, Extended (Open Response)** A box for typing text is provided; valid characters and maximum amount of characters can be defined. Can include the option of one or more of the following: cut, copy, paste, undo, and redo buttons; a progress bar for indicating usage of the total amount of available space; character count; line count; word count; and spell checker.
Matching	▪ **Match (Drag-and-drop), Category or Simple Drag and Drop** Test-taker drags individual elements into categories. ▪ **Match (Drag-and-drop), Single aka Gap Match** Test-taker indicates 1-to-1 matching relationships between elements. Text-based content only.
Hotspot	▪ **Object Selection, Graphic Change, Single** Test-taker selects one image or portion of an image. ▪ **Object Selection, Graphic Change, Multiple** Test-taker selects multiple images or portions of an image.
Position Object	▪ **Icon Placement, Coordinate-based** Test-taker places a single icon; response is returned as x/y coordinates; correct answer is defined as a set of coordinates
Inline Choice	▪ **Inline Choice (Cloze), Single** A drop-down menu is used below an image or a range of text to choose a word or phrase from several options
Hot Text	▪ **Hot Text Interaction** Test-taker selects one or more runs of text embedded within a larger context, such as a reading passage.

QTI Interaction Types Supported by TestNav	
Custom	▪ **Graph, Line** Test-taker plots points which are connected by line segments in order of ascending x-values ▪ **Graph, Scatter-plot** Test-taker plots points on a grid ▪ **Graph, Polygon** Test-taker can plot points, which are connected by line segments in order in which the points are plotted, with the option for a "close the shape" button that connects the first point to the last and fills in the polygon with a color ▪ **Manipulative** Manipulatives are objects that a student can change via animation, simulation, or some other method ▪ **Vertical Bar Graph** Test-taker clicks and drags on bars ▪ **Text Extraction** Test-taker can drag text and drop into an element ▪ **Text Highlighter** Test-taker can click on text to highlight ▪ **Interactive Number Line** Rays and segments can be plotted on a number line using a specialized interface for activating a segment for editing or deleting that segment ▪ **Zoom Number Line** Provides a zoomed in view to a number line for finer-grained detail ▪ **Fill-In Blanks with Equation Editor** The test-taker can use the equation editor to fill in blanks within a mathematical expression ▪ **Constructed Response with Equation Editor** The test-taker may create a mathematical expression using the equation editor ▪ **Function Graph** The test-taker chooses a graph shape and then manipulates two points to graph a function ▪ **Solution Set Graph** The test-taker can create two lines, choose a dashed or solid style for each line, and shade one or more areas created by the lines to indicate a solution set for two inequalities ▪ **Composite Graph** Two sets of labeled and color-coded elements can be added to a single coordinate grid using the functionality of scatter plot, line graph, coordinate grid graphing, and polygon graphing in combination with one another

The TestNav platform provides item types that are consistent with IMS Global QTI specifications.

TestNav also offers usability features that are made available at the time the test is created. The following figure identifies some of these features. Like tools and item types, this list will continue to expand.

Additional Usability Features	
Feature	**Description**
Exhibit	Allows for use of a formula sheet, images, periodic table of elements, etc.
Copy and Paste	Copy and paste functionality can be enabled for text editor items.
Pop-up Glossary	Provides a definition of specific words.
Video Passage	Allows for the use of a specific video to be used within an item.
Audio with Volume Control	Enables an item that requires audio to be made available with volume controls.
Passage Paging	For long reading passages, provides a means to include a forward and backward navigation buttons on the passage.
Timed Test	Enforce a time limit on a test session.
Passage Scrolling	The student can view long reading passages by scrolling.
Test Sections	Items in a test can be grouped.

Additional Features: Pearson will continue to build out features in TestNav that align with industry standards and best practices.

TestNav also contains the following test-taking tools that can be configured for each individual item.

TestNav Tools	
Tool	**Detail**
Calculators	Texas Instruments graphing, scientific, basic Generic 4-function, 5-function (includes square root), scientific
Ruler	Inches, Centimeters, 1/4" increments, 1/8" increments
Mathematical Compass	Draws circles and arcs of a specified radius
Protractor	Measures angles
Answer Elimination	Crosses out (or restores) a distractor

TestNav Tool Set. Pearson continues to improve its online tools to keep pace with interfaces that students already find familiar.

Built to Meet the Demands of Next Generation Assessment

A successful testing experience for Arizona students depends on test delivery technology that is dependable, consistent, and secure from data interception and cheating. The test delivery system must also be scalable, and the implementation must not overburden school technology staff.

When evaluating an online test delivery system, Arizona should consider five basic criteria to determine which system is the right fit for its assessment program:

- Security
- Consistency
- Scalability

- Reliability
- Cost of Ownership

Test Delivery System	
Common Online Testing Requirements	**Pearson Solution**
Security	
Test content must be secure from interception during transfer and storage.	TestNav test data are encrypted on a Pearson server and remain encrypted in transit and on the caching server. TestNav decrypts the test items only in the temporary memory of the student's testing computer. Any Arizona data intercepted would be completely indecipherable.
The interface must not enable students to cheat in any way. The test delivery system must prevent saving, sending, or receiving any data while testing.	TestNav puts the testing computer on "lockdown," also known as "kiosk mode." No functions other than testing are enabled. TestNav is engineered to prevent access to the following: ▪ The desktop ▪ Cut, copy, and paste functions ▪ Any other applications ▪ Messaging and email ▪ Printers ▪ Screen capture
The delivery system must incorporate strong protocols for authenticating the user and protecting user and test data.	To verify that the right Arizona student is taking the right test, TestNav requires the correct student ID, combined with a unique test code. These login credentials are used to authenticate that the right student is taking the right test and will work only after the proctor approves the student and starts the test session. In addition to strong authentication protocols, TestNav maintains the examinee state external to the client. This means that the only data on the client machine at any one time are the current item. Other information about the student and the test is maintained on the server.
No data shall remain at the testing site after the student finishes the session.	Test items are decrypted only in the temporary memory of the testing computer and never written to the clipboard or hard drive. When TestNav terminates, the memory is flushed and no trace remains of the assessment or the student's answers.
Consistency	
To keep entry costs low, the delivery system must operate on older computers and operating systems while maintaining a comparable experience for users.	TestNav is designed to run on the computers, operating systems, and browsers already found in most Arizona schools. This includes many systems that are near the end of their service life. With browser-based TestNav, no additional hardware or software is typically needed. The user interface and test content are delivered through the same web browser that students use every day.
The system must function responsively while delivering the items specific to Arizona.	The file size of a test item affects technology requirements. For example, items containing video can place greater demands on local infrastructure. We will work with the ADE and Arizona schools and districts to match the test to the overall computing environment.

Test Delivery System	
Common Online Testing Requirements	**Pearson Solution**
The user experience must be visually consistent across approved computers and operating systems.	TestNav will present a consistent user experience. Pearson has designed a web application using responsive design principles. This means the interface scales itself based on the size of the device's screen. Users will have the same online tools, and the presentation will have the same look and feel, adjusted to the device used to deliver the test.
Scalability	
Servers must be able to balance peak loads and deliver consistent response times.	Pearson emphasizes scalability in both the server and client environments. We will examine the following three areas to avoid assessment bottlenecks: ▪ Types of computer used for testing ▪ Bandwidth necessary for the specific test ▪ Server capacity and balance Pearson uses a variety of techniques for configuring and balancing servers for consistently quick response time. In addition to traditional load balancing and hardware deployment techniques, we use software technologies that provide clustering capabilities.
The vendor must constantly test and verify system capacity.	We will conduct regular performance testing to simulate how the entire Arizona environment will actually behave on test day. Networks change and computer usage patterns change, so we observe and anticipate the needs of each test administration. Pearson has successfully delivered more than 429,000 tests in a single day.
The solution must address school bandwidth issues.	TestNav includes lightweight, secure local caching software. Proctor caching enables schools and districts to download test content only once and deliver it securely from local servers, instead of relying on potentially shaky or insufficient Internet connectivity for content delivery during testing. If a school or district has the necessary bandwidth or existing caching software, they can test without proctor caching.
Reliability	
The delivery system must save student responses in case of power or network failure.	TestNav is designed to store all responses that students have made, even when a power or Internet disruption causes delay or failure. An early warning system will minimize the effect technical failures have on the student's Arizona assessment experience.
The system must provide the location of saved responses.	If a student's workstation is unable to transmit responses to the testing server during a test, the early warning system saves the responses to an encrypted backup file in a configurable location. This activity occurs in the background while TestNav continues to provide test questions to the student, who may either continue testing or exit without losing data.
Testing must resume with minimal disruption.	TestNav and the early warning system are designed to enable a student to resume the test after an interruption at the same point in the test where the interruption occurred. When the network connection resumes, TestNav uploads the saved responses to the testing server and erases the encrypted response file automatically.

Test Delivery System	
Common Online Testing Requirements	**Pearson Solution**
Cost of Ownership	
The client software shall require no downloading or installation.	TestNav requires only a standard computer, operating system, and browser to deliver functionality when using desktops or laptop computers. Solutions that require special software downloads to be installed on each computer place a heavy burden on local technology support staff, first to initially install and configure the system and then to re-touch each machine when there are updates or patches. These steps are completely unnecessary with TestNav.
Any software maintenance must be vendor's burden.	TestNav resides on Pearson servers, where we maintain it. The software will never need to be maintained on Arizona desktop or laptop computers.

Evaluating an Online System. When evaluating an online test delivery system, five basic criteria—security, consistency, scalability, reliability, and cost of ownership—should be considered to determine which system is the right fit for the Arizona assessments.

System Scalability

With TestNav, Arizona's assessment program will have the capacity to allow smooth functionality. As a leader in large-scale, high-stakes assessments, Pearson has the experience and infrastructure to provide the capacity necessary for statewide usage within relatively narrow testing windows. Vendors that lack our experience and resources often struggle to quickly respond to changing usage.

Large-scale, high-stakes testing requires large-scale, high-stakes planning, capacity, and monitoring. Pearson will study the capacity of the Arizona assessments program and how its demands interact with our existing systems. Our software is designed to scale using traditional load-balancing through hardware-deployment techniques and the capability to use clustering.

In addition to technologies that can handle the load, the ADE needs a team with proven methods for assessing usage patterns and the experience to make the correct decisions. Pearson monitors and plans for system load, and we proactively scale the system as needed to meet peak usage needs. We conduct regular performance tests that simulate true system usage patterns, using a range of automated testing tools and load generators. These systematic evaluations enable us to plan for the impact of each client's programs.

Pearson's flexible, dependable system means Arizona schools and districts will have the capacity to successfully complete testing on time. By monitoring our large multi-server system under high loads, we are able to determine which system resources are affected so we can extrapolate the capacity of various configurations and situations. These steps enable us to evaluate the impact of large numbers of users accessing the system simultaneously so we can scale our capacity accordingly.

Vigilant Applications and System Monitoring

Our comprehensive application and system performance monitoring tools and procedures include the following:

- Automated systemic monitoring performed by our data center host and Pearson, 24 hours a day, seven days a week—including the operating system, central processing unit and memory utilization, and network health
- Automated database-level monitoring by Pearson with proprietary monitoring tools that provide early warnings on performance and capacity thresholds
- Automated external monitoring from geographically dispersed locations performed 24 hours a day, seven days a week measuring performance and availability from a client perspective
- End-to-end diagnostic monitoring providing performance measurement of application and database components

ACT Aspire

Computer-Based Administration

Using Familiar Technology Available in Schools

ACT Aspire summative and interim online assessments for grades 3-10 will be supported using technology systems developed by Pearson, including the ACT Aspire administrative management system, and Pearson TestNav online test delivery system. The ACT Aspire platform will run on Macintosh and PC platforms.

[redacted]

Stronger Technologies

Arizona will benefit from Pearson's vast experience in online administrative and test delivery technologies, including the following:

- Computer-based test administration with paper and pencil as an option
- Browser-based "zero footprint" student interface when using desktops or laptops
- Robust online reporting for administrators, students, and parents to facilitate timely modifications to student learning and intervention strategies
- Robust capacity with advanced system scalability and contingency provisions if Internet inconsistencies threaten the test window

One Secure, Reliable Website for Managing the Arizona Assessments

Arizona educators can use a single secure interface that simplifies SBCA administration tasks. Newly developed from years of experience, the ACT Aspire platform has been designed as a reliable, functional, and flexible way to manage assessment data. Hierarchical role-based access will protect Arizona data.

The ACT Aspire Data Management Platform

The ACT Aspire technology platform for Arizona is the result of years of evolutionary development in the assessment delivery field, development that continues with rapid responses to changing needs. We improve the ACT Aspire platform in both major releases and incremental improvements within a rapid development cycle. A static solution does not have the flexibility and functionality that the ADE requires for the Arizona assessments.

Arizona districts and schools will use one comprehensive ACT Aspire platform site to accomplish the following:

- View and edit student data as loaded or entered during the precode process
- Enter student data prior to or during the testing window
- Maintain student and test data
- Complete information to indicate how results are returned to school districts
- Assign appropriate form distribution plans
- Schedule online student testing, including providing necessary student login information
- Specify student accommodations
- View test information as defined by user access

The administrative system used by the Arizona assessments will protect student privacy with industry-standard security measures. The ACT Aspire platform employs Hypertext Transfer Protocol Secure (HTTPS) to provide end-to-end encryption, and user permissions are configurable. Districts, schools, and the ADE can control data access appropriate to the task through a rules-based system of user roles.

ACT QualityCore

CBT Administration (2014–2015 Administration)

The ACT QualityCore CBT program is currently offered as multiple choice tests for all 12 courses as well as multiple choice and constructed response tests for English 9–12. ACT currently offers CBT administration for both multiple choice and constructed response only for English courses. ACT is committed to offering computer-based testing for all courses in the winter of 2014–15. The administration of ACT QualityCore for Arizona's end of course assessments will be available to users in Arizona on this standard platform.

Assessments include problem-based questions embedded in both academic and real-world contexts that are accessible and relevant to high school students. These real-world problems require practical applications of concepts, theories, principles, and process.

Flexible options include:

- Two 45-minute test sessions—multiple choice and constructed response, or multiple choice only
- Computer-based or paper-and-pencil format
- Standards-based reporting

Results support informed decisions: Each assessment will report a total score and three to five diagnostic subscores associated with the ACT Course Standards. These subscores enable educators to drill down to specific areas of strength and weakness in student learning.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Custom Assessment

Computer-Based Assessments

For custom assessments built to Arizona specifications, administrative operations would be managed through PearsonAccess, the same system currently used for the AIMS and AZELLA, and test delivery would be provided through TestNav. Arizona schools and districts participating in the spring 2014 PARCC online field test will have the opportunity to experience online testing through PearsonAccess and TestNav firsthand, offering a distinct advantage for alleviating program startup risks and training workloads if this same combination were used for custom assessments to support Arizona's first operational SBCA administration in spring 2015. Based on Arizona's experience through AIMS and AZELLA, and upcoming experience through PARCC, additional Arizona-specific features, functionality, or configuration options such as local online score entry could also be requested for custom SBCA assessments.

Requirement

b. If applicable, describe the technological specifications for the administration of the assessments. This should include specifications for computer hardware, input devices, security requirements, bandwidth, web browser requirements, and platform software.

Response

Infrastructure for Online Testing

Our TestNav system is compatible with a wide range of operating systems and browsers. We routinely monitor and evaluate the minimum and maximum hardware and software releases recommended for use. As third-party hardware and software vendors release new versions, we update our system requirements for online test delivery accordingly.

By supporting workstations commonly found in schools, Arizona can use computer technology already in place statewide.

The list of supported software versions evolves appropriately as new operating systems and browsers are released. As of August 2013, the TestNav system supports Windows (XP, Vista, 7, 8), Macintosh (10.5, 10.6, 10.7, 10.8), and iOS6 with Microsoft Internet Explorer 9, 10, Chrome, Firefox and Safari browsers. By fall 2014, Chromebook, Android and Linux operating systems will also be supported.

Arizona schools will also have the option of using touch-screen devices for high-stakes test delivery. We developed the next major release of the TestNav platform for tablets first and then redesigned its desktop-computer functionality to be comparable with the touch-screen environment.

With every release of the TestNav system, we test using the most current list of supported devices, operating systems, and browsers to be confident that our customers will be ready on test day. As new devices and software become available, we carefully evaluate their technology and usability for a secure, high-stakes assessment environment before adding them to our supported list.

Proctor Caching

With TestNav, Arizona districts and schools will have a highly functional, reliable, and responsive testing environment for delivering online Arizona assessments with existing local infrastructure. To maintain responsiveness and provide additional local flexibility as tests become more media-rich, Pearson provides local proctor caching software. For schools with concerns about local bandwidth or internet connectivity issues, we recommend proctor caching test content.

Pearson proctor caching is a simple, secure option that enables the TestNav system to deliver interactive tests using limited Internet bandwidth. With proctor caching, a test administrator downloads test content only once from the Pearson server to the local caches. Safely encrypted, it resides on a local computer and is delivered during testing to each student's computer, where the TestNav system decrypts and displays it. The local Internet service provider usage is minimized when using proctor caching, therefore reducing the potential for Internet delays.

Flexible Delivery Options

Our proctor caching server is not a one-size-fits-all solution. While some systems require a separate content caching server in each school or fail to provide adequate tools for optimizing local test delivery at all, our solution can be deployed on existing hardware at a scale that best fits local needs and capacity at the lab, school, and district levels.

The TestNav platform does not require schools to cache content, and using additional tools that we provide, local personnel can determine whether proctor caching is necessary. Many larger districts already have a caching server in place to meet their Internet needs, and Pearson does not require that they use our software. We do, however, highly recommend the use of proctor caching for best online testing performance, particularly in bandwidth-challenged schools.

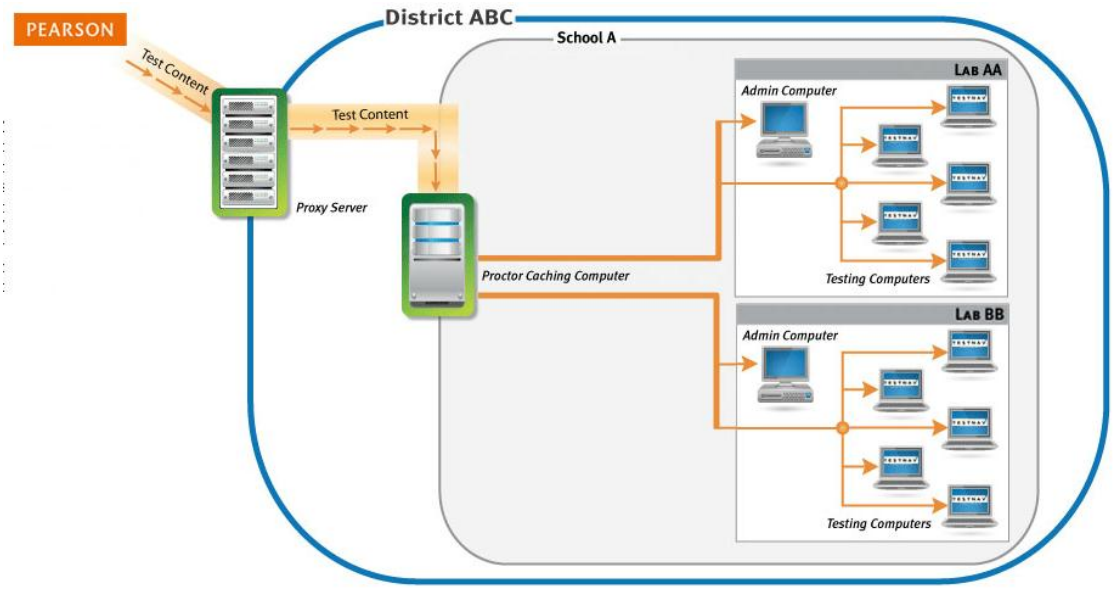
To predict user capacity, local technology personnel will use the Pearson System Check for TestNav. If the planned number of simultaneous testers would cause bottlenecks, then the local technology department can choose proctor caching and configure it to fit their specific needs. The system check tool

evaluates connectivity to the proctor caching workstation in the school or district and can help determine the optimal number and locations of caching computers needed.

Some states prefer to mandate proctor caching for all districts based on the size and complexities of tests that are being delivered. A district that normally has excess Internet bandwidth may decide that proctor caching is an unnecessary step. Conversely, a district struggling with capacity could find that caching content is essential to providing an uninterrupted testing experience.

High-Capacity Server Not Required

The ADE and Arizona districts will not need to purchase server-class machines for proctor caching. Nearly any computer commonly used to access the Internet in schools is sufficient to cache and serve media-rich content to student testing computers.



Increasing Test Administration Speed. With the proctor caching feature, Arizona districts and schools can eliminate multiple data streams during online test administrations, increasing speed, protecting students from disruptions in internet connectivity, and reducing bandwidth demands.

Early Warning System

When an Arizona student enters a response, the TestNav Early Warning System saves the response even if network traffic is interrupted. Pearson takes each student response seriously, and we have rigorously designed our systems to preserve data with a high degree of fault tolerance. The TestNav system tolerates Internet and local network interruptions, provides advance warning of possible problems, and encrypts responses in a backup file that is routinely refreshed in the background. As a further precaution, when the early warning system detects a potential problem with the designated save location, instructions appear on-screen for the test monitor to create a new location to save responses.

Our test delivery system is also designed to allow a student's test to be resumed after an interruption—at the same point in the test at which it was interrupted. In our standard testing environment, the TestNav

system sends student responses to the Pearson testing server in real time, while the student is still testing. If a disruption in communications with the server or local content cache lasts more than a few seconds, the Early Warning System saves the encrypted response to a backup file and provides the student with a message that their work has been saved.

At that point, depending on the circumstances, a student can either attempt to continue testing or exit the system and resume testing later with their saved data, including any annotations, highlights, or other intermediary work they may have completed while working on an item.

In many cases, when local network connectivity is maintained and only connectivity to the Internet is lost, students are able to continue testing from the secure local testing cache without any awareness of outside connectivity issues. In such instances, when the student reaches the end of the test, their work is saved locally until connectivity to the TestNav server is restored. When the network connection is restored and the proctor has reactivated the student to resume testing, the TestNav system automatically uploads the student's saved responses to the testing server and then erases the encrypted response file. In this way, the TestNav system shields the student from the connectivity problem, presenting items and saving the student's answers with no interruption in testing.

System Check Tools

Arizona schools and Pearson systems will be ready to deliver assessments online and on schedule. District technology personnel can use the Pearson readiness toolset for quantifying bandwidth and planning allocations. When questions arise, they can contact Pearson technical support for assistance. Readiness for a large-scale assessment depends on Arizona district personnel getting accurate information and TestNav-specific tools to analyze systems and assist in making any necessary corrections.

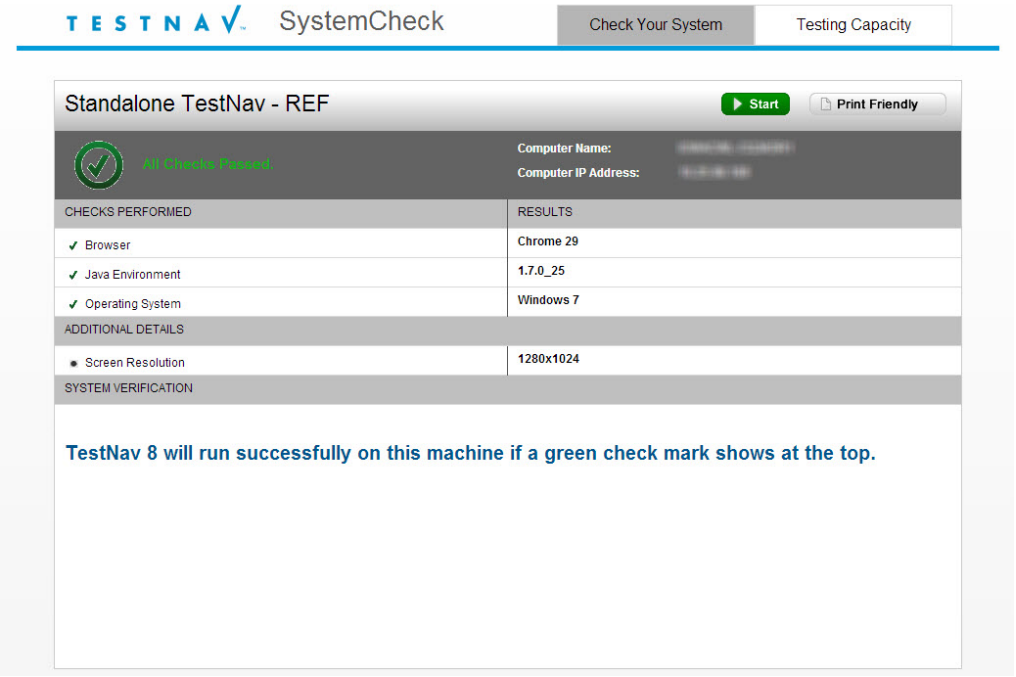
Arizona district technology personnel can check whether schools and district infrastructure meets minimum requirements for online testing, including client hardware, connectivity and bandwidth, proxy servers, and firewall configuration. Our toolset quantifies the environment at a point in time and estimates how many tests the school or district can deliver concurrently. Without this live calculation, capacity planning would be based on a guess.

Using the Pearson toolset, districts can quantify bandwidth and plan allocations during testing periods. System Check for the TestNav system is available year-round, and district technology personnel can consult with Pearson technical support staff.

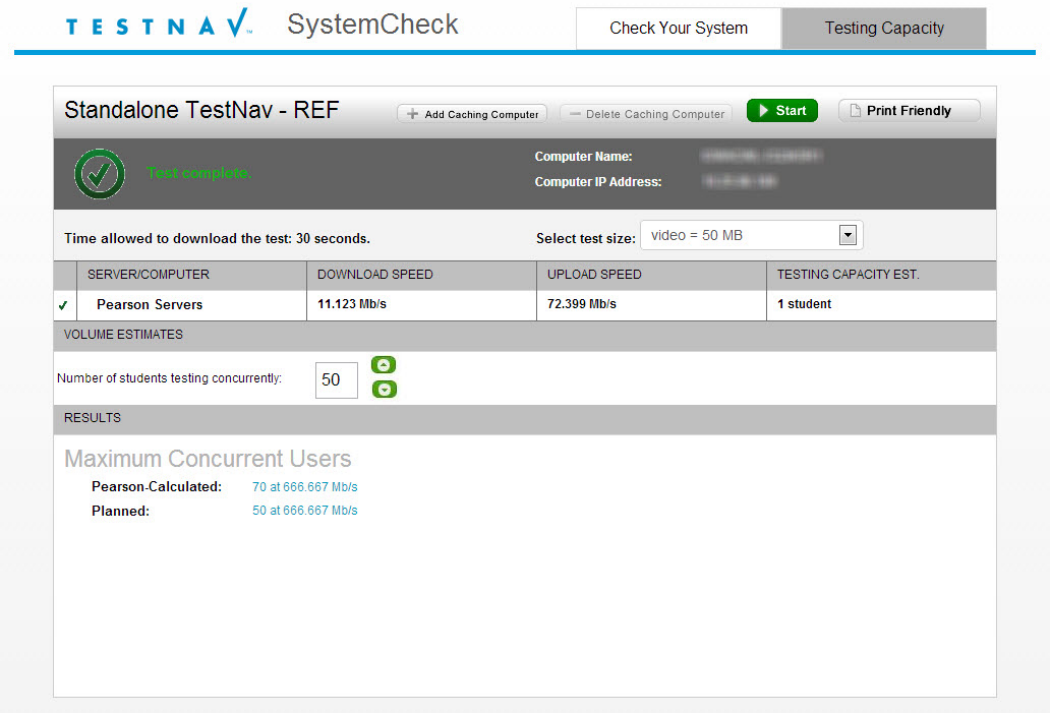
Our system check tool includes two components:

- **Check Your System** verifies that the computer meets the minimum requirements for the TestNav system.
- **Testing Capacity** checks connectivity to Pearson servers and any proctor caching software and allows users to add or delete a proctor caching workstation as warranted.

Over the coming weeks and months, local Arizona technology support personnel will be trained in using these tools as part of their preparation for the upcoming PARCC online field tests.



Check Your System. The Check Your System page determines whether the computer meets the minimum requirements.



Testing Capacity. The Testing Capacity page checks connectivity to Pearson servers and local proctor caching servers, helping to identify testing capacity at a specific point in time. Users can add or delete a proctor caching workstation as warranted.

Technology readiness is critical requirement for Arizona districts and schools to successfully implement online testing. In addressing technology readiness issues, Pearson will draw on our experience as the contractor for the SBAC/PARCC Technology Readiness Tool (TRT). Under the contract, Pearson developed and deployed an open-source, online tool to evaluate readiness of states, districts/LEAs, and local schools to implement new online tests. The TRT allows local schools, districts, and LEAs to capture and report key readiness indicators, including the following:

- Number and type of computers
- Local network and bandwidth infrastructure
- Local staff resources
- Other information needed to evaluate overall technology readiness to support the transition to online testing

In addition to developing the tool, Pearson's scope of work includes hosting and administering the TRT through a network of state readiness coordinators (SRCs), and producing school, district/LEA, and state technology readiness reports.

ACT QualityCore

The following technical specifications detail ACT's currently recommended hardware and software requirements to support ACT QualityCore.

ACT QualityCore Technical Specifications	
Student Testing	A secure browser is required for student testing. Chromebook is NOT supported for student testing.
Linux	Any distribution that supports Firefox 22.0, Chrome 28
PC Operating System	Microsoft Windows® XP (SP3), Vista, Windows 7 & 8
PC Web Browser	IE® 9.0 or Firefox® 20 or higher, Chrome 27 or currently supported versions
Mac Operating System	MAC OS X 10.6 or higher
Mac Web Browser	Firefox® 16.0, Safari 5.05, Chrome 27 or currently supported versions
Processor	Processor speed should meet or exceed the minimum required by the user's operating system and browser
Memory (RAM)	Available memory should meet or exceed the minimum required by the user's operating system and browser
Bandwidth	30 Kbps of dedicated bandwidth to the Internet and back for EACH workstation is recommended. Note: Many factors can impact the online test session. For consistent testing experience, ACT recommends network support such that no item takes longer than 2-3 seconds to load
Resolution	1024x768 resolution or better. Smaller screens sizes and nonstandard screen resolutions

ACT QualityCore Technical Specifications	
	may require additional user scrolling or restrict the readability of text
Pop-up blocker	Browser set to allow all pop-ups.
Workstation cookies	Cookies should be set to “accept all cookies.”
Workstation cache	Cache should be set to the lowest possible setting that can be used for your specific computer or network
Network proxy, Firewall, Security, Content filters	Set to bypass the specific domains and IP addresses. TCP/Proxy servers and firewalls configured to allow unrestricted http on port 80 and https on port 443 to the Internet and also to permit response traffic
Network Proxy and Authentication Settings	The session time-out value should be set to be at least equal to the length of the test or a minimum of 60 minutes
JavaScript and Active Controls	Browsers must have JavaScript and Active Controls enabled

[redacted]

[redacted]

[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Looking Forward

Pearson is committed to continue moving forward to support and use the latest technology to improve the testing experience for students, administrators, and educators. Through the years, we have continued to invest in our systems and platforms and have guided their evolution as we discover what works well and where there is room for improvement. With each new generation of our software, we have adjusted and improved our designs for efficacy and efficiency. We have made changes to support new technologies that allow for greater flexibility and enhanced features. User interfaces have been adjusted to better suit the needs of system users, making screens and workflows easier to understand and faster to use.

We are continuing these steps forward now by developing the capability for students to access delivery systems through standard browsers on the same computers they use every day for instruction and low-stakes testing (native applications will be used for mobile devices). Most high-stakes assessment delivery systems require school personnel to install a proprietary browser, while others have minimum requirements that are not fully verified. Pearson's delivery system is designed to take advantage of a wide range of hardware, software, network resources, and technology infrastructure that exist at the local districts and schools.

These continued developments will eventually benefit all computer-based testing we offer. These accomplishments are just some of the ways Pearson continues to advance our capabilities to provide the best solution for our customers.

Requirement

c. Describe any available training and technical support that is available for the computer-based assessments.

Response

ACT Aspire

Customer support will be paramount to implementation and administration of the ACT Aspire assessments. In addition to the ability of administrators and teachers to email or call our toll-free customer support team, we will provide comprehensive trainings as follows:

- Overview of ACT Aspire
- CBT system readiness (see below)
- Data management (loading organization and student data)
- Order management (ordering materials and optional services)
- Test setup and management (creating, viewing, editing, and monitoring test sessions)
- Test administration (printing authorization tickets and launching test sessions)
- Post-test interpretation and support (accessing reports and interpretation materials)

Readiness Support

For computer-based testers, customers have access to readiness support through the following tools and trainings:

- A system overview that outlines technical aspects of TestNav
- An overview of the TestNav Early Warning System (EWS), which describes the integrated functionality that gives TestNav a high degree of fault tolerance and provides additional fail-safes in the event of unexpected network disruptions during online testing
- A system check tool that helps district and school technology personnel better assess the overall readiness of their local network for online testing
- Proctor caching, a Pearson-supplied software used in conjunction with TestNav to reduce bandwidth requirements for school or district networks and accelerate the delivery of test content. Proctor caching enhances the performance of existing technology in schools. No additional hardware is required

We can also provide onsite support as needed throughout this readiness phase; however, the scope and nature of such requests may incur additional charges that can be borne by or shared with the district(s) or school(s) involved.

Phone Support

Providing Responsive Customer Service to Arizona

To provide call center service to handle phone calls and emails regarding the SBCA, Pearson offers a Customer Service Center (CSC) with a strong history of support for diverse state and national assessment testing customers, including both the AIMS and AZELLA. Formed in 2003, our CSC services approximately 110,000 contacts per year and currently supports 58 individual assessment programs and six unique platforms/products across multiple states.

Our CSC has experienced staff members available to respond to the needs Arizona District Test Coordinators. Arizona callers will receive prompt and personal service from qualified staff in sufficient numbers. Our team members are experienced and competent customer service professionals who can answer questions efficiently and thoroughly. To date in 2013, customers who completed the customer satisfaction survey indicated the CSC met or exceeded their expectations 95 percent of the time.

The Pearson CSC offers a range of services to effectively support callers from Arizona:

- Callers will have **access to live support** via email, chat, and toll-free phone from 7:30 a.m. to 4:30 p.m. (MST) each weekday, excluding federal holidays.
- **Specialists** with detailed program information and focused expertise in the database, network, infrastructure, and software components of Pearson's web-based services will respond to questions that cannot be resolved at the initial point of contact.
- Using **incoming call routing to manage call flow and expedite service**, we direct calls to team members best able to answer the question, including specialists to whom calls may be escalated. We

also offer **email support, including automated return email** to confirm receipt. For email received during off hours, we will promptly respond during the next regular business day.

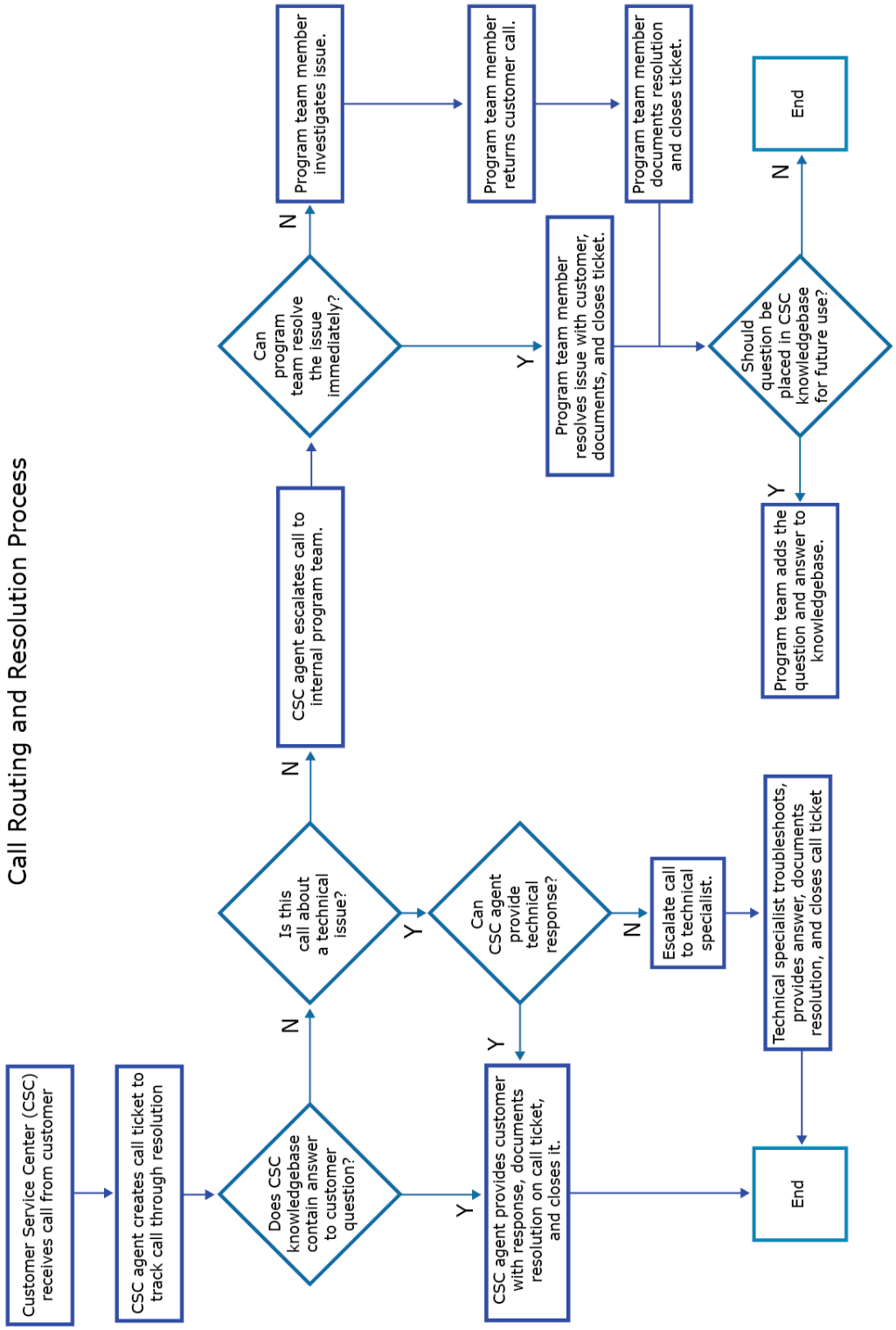
- Our solution provides **access to the same knowledge base** used by CSC staff to obtain information needed to answer questions and provide information. The knowledge base is continuously updated in collaboration with the program team and available for use by callers.

Scalable Capacity for Efficient and Dependable Service

Our flexible, scalable operations allow us to provide consistent, quality service during peak periods by routing calls as needed. Operating across a network of centers in Iowa, Minnesota, and Texas, our CSC is more efficient because it is independent of a single location.

We create a contingency plan for peak periods using scheduling and forecasting information from the program team and CSC data. In periods of anticipated peak call volumes, we add trained personnel who can log in to the phone system and begin taking calls. This built-in additional capacity will reduce wait time for Arizona callers during peak volumes.

The following figure shows our call routing and resolution process.



Call Routing and Resolution. Our escalation protocol will efficiently route Arizona callers to the appropriate personnel.

Trained, Effective Staff

To meet Arizona's needs, we provide qualified and well-trained staff. CSC personnel are trained to quickly determine the scope and impact of reported problems and efficiently route and resolve them. To provide the high quality service for the SBCA, training for new CSC staff would include the following:

- Overview of knowledge, procedures and processes used for supporting our customers
- Pairing up with an experienced staff member to begin listening to customer calls and receiving instruction on phone system operation
- Detailed introduction to and instructions for use of specific websites, products and platforms
- Familiarization with tools used to house knowledge and call information
- Role playing with other staff members and practice with specific call scenarios that may be encountered
- Live customer responses with mentoring by a staff member

Staff members receive regular feedback from the supervisory staff and also receive individual call statistics, so they can see their own progress. Arizona test coordinators benefit by having access to people who can respond appropriately even when they are new to their job.

Customer Service Systems and Tools

The Pearson CSC offers the following systems and tools to effectively serve Arizona callers:

- **Telephony.** By effectively creating options from which the caller chooses, we use the Avaya telephony system (Avaya G3R Version: V15 w/CM 5.2) to route calls to CSC members trained to handle specific types of issues. This routing allows for a more efficient flow of calls to CSC staff who have the best opportunity to resolve issues during the first call.
- **Chat.** The CSC team uses LivePerson to provide real time customer support. This allows customers the flexibility to interact with agents online.
- **Knowledge Base.** The CSC team uses the Servigistics knowledge base (Kaidara Advisor 4.4) to access the information needed to answer Arizona caller questions allowing for consistent responses across team members. Because our knowledge base is continuously updated in collaboration with the program team and, as needed, approved by the ADE information that may change is promptly available to CSC staff for use in answering questions.
- **Incident Tracking.** Contacts are tracked by the HP Service Manager Client 7.11.259, a secure incident-tracking software tool tailored specifically for the Pearson CSC. With each incoming contact, CSC staff generates a ticket that enables tracking the incident through to resolution. Tickets are archived and ticket numbers are also provided to the caller. Tickets may be escalated to specialists for their documentation when handling calls to generate a full history of the issue.

- **Quality Monitoring.** We use the Verint Impact 360 Work Force Optimization Suite (v. 11) to provide high-quality service to Arizona. Through recording/monitoring of calls, CSC team members receive regular feedback from CSC supervisory staff and also receive individual call statistics, so they can see their own progress.
- **Caller Satisfaction Surveys.** To confirm we are meeting your needs, we email surveys to gain valuable information on how Arizona callers view their experience with CSC staff and our response to your concerns.

ACT QualityCore

The ACT QualityCore program has a number of administration manuals available for training and technical support. Through its more than 50 years delivering standardized assessments, ACT has designed, developed, and implemented countless test supervisor manuals informed not only by program requirements, but by literally tens of millions of testing experiences in schools across the nation and around the world.

ACT QualityCore has been designed to operate effectively “off-the-shelf” or configured for delivery on a contract-specific basis. The program includes comprehensive test administration manuals that would guide district- and school-level administrators through both the paper-and-pencil and computer-based modes of testing. ACT has also participated in developing supplements for specific contract customers for the ACT QualityCore program.

ACT QualityCore test administration manuals include:

- Online Users Guide
- Test Supervisor’s Manual
- Room Supervisor’s Manual
- Instructions for Special Testing
- CBT Room Supervisor’s Manual
- Instructions for Accommodated Testing
- Technical Manual

[redacted]

Requirement

d. If the assessment is computer-based is there a paper / pencil option?

Response

Although the benefits of next generation assessment are maximized through the enhanced capability that online administration affords, reality dictates that not all schools and districts will be able to fully support online testing at once. Our experience helping other states through this transition has shown that even with the strongest state support, it will typically take anywhere from three to five years before online testing can be considered fully implemented, although these timeframes are slowly shrinking as schools' use of technology grows over time. Even then, paper accommodations for testing are still needed for some students or in other isolated instances.

ACT Aspire

Solutions for Transitioning Schools

Arizona may need to continue paper and pencil administration of the SBCA while transitioning to online delivery. As proven through our support for the AIMS and AZELLA. We can easily accommodate this option. In addition to online delivery of the ACT Aspire assessments, we also offer the option of administering the assessments in paper/pencil format.

Our approach for paper includes pre-printing student information on answer documents prior to distribution, thus taking the burden off of educators to hand grid answer documents for each student or apply pre-ID labels. This presumes that the ADE or Arizona schools have submitted data to ACT Aspire through our online student-data upload process.

If Arizona's student data submission window expires and additional students need to take the paper assessments, we offer an alternate process to enable educators to submit files for those additional students. Depending on the level of urgency, we can create pre-printed answer documents and ship them to the schools along with test booklets, or educators can submit student data files through the ACT Aspire portal and receive a system-supplied identification number for each student. This number can then be used to grid on a blank answer document supplied via the overage shipment. The system identification number serves as the link between the student's answer document and his/her demographic data in the portal.

We will ship paper materials to Arizona schools using a carrier with the capability to track shipments to the location specified within an organizational data file submitted by the ADE to the ACT Aspire team through our online system.

ACT QualityCore and The ACT

As previously described, paper and pencil options are also available for ACT QualityCore and The ACT.

Custom Assessments

As previously described, Pearson's test administration, processing, scoring, and reporting systems were specifically designed to provide greater flexibility and reduce administrative effort by facilitating the transition from paper-based to online testing at a pace that fits our customers' capabilities and preferences.

Custom assessments developed for the SBCA would be equally supported in both paper and online modes through PearsonAccess for as long as needed. Our history of research in paper and online comparability, coupled with our experience managing assessments in parallel administration modes, would enable the ADE to maintain a high-quality, defensible assessment system. By regularly monitoring district and school capacity and readiness for online testing and carefully tracking historical paper and online use rates, Pearson would support the ADE in annually adjusting its transition strategies, and plans for future test development and administration, helping to more effectively project and manage ongoing costs.

4. Assessment Administration

Requirement

a. Describe the total anticipated testing time for each assessment (mathematics, reading, and writing) by grade level. If computer-based, include the calculated student to device ratio.

Response

ACT Aspire

Testing Times

The figure below shows the total anticipated testing time for each ACT Aspire assessment. These numbers are standard anticipated testing times; if a student requires extended time as an accommodation, these can be extended based on the student's needs.

ACT Aspire Anticipated Testing Times					
Grade	**English**	**Writing**	**Reading**	**Math**	**Science**
3	30	30	60	55	55
4	30	30	60	55	55
5	30	30	60	55	55
6	35	30	60	60	55
7	35	30	60	60	55
8	35	30	60	65	55
9	40	30	60	65	55
10	40	30	60	65	55

ACT QualityCore

The ACT QualityCore End-of-Course (EOC) Assessment consists of either two multiple-choice parts or a multiple-choice part and a constructed-response part. When the multiple-choice and constructed-response parts are used, the multiple-choice part should be administered first, followed by the constructed-response part. Any day during the testing window can be selected to administer the test. All tests in a given course must be administered on the same day(s) within the testing window. Each section of the EOC assessment requires 45 minutes of testing time. Sections of the test can be administered in one session or two separate sessions. Most tests will be administered to students in the classroom during

the regular class period. However, students in a given course may be tested outside of the regular class periods if facilities and schedules permit.

The ACT

The ACT includes tests in English, mathematics, reading, and science. Test content and times are as follows:

- The English test (45 minutes) includes a total of 75 test items covering usage and mechanics skills in the areas of punctuation, grammar and usage, and sentence structure. The English test also includes rhetorical skills in the areas of strategy, organization, and style
- The mathematics test (60 minutes) includes 60 test items in pre-algebra, elementary algebra, intermediate algebra, coordinate geometry, plane geometry, and trigonometry
- The reading test (35 minutes) includes 40 test items in the areas of prose fiction, humanities, social studies, and natural sciences
- The science test (35 minutes) includes 40 test items and covers topics in biology, earth/space sciences, chemistry, and physics in formats including data representation, research summaries, and conflicting viewpoints
- The 30-minute writing test is an optional part of The ACT

Custom Assessments

By constructing test blueprints specific to Arizona's own needs, the ADE could create assessments as long or short as desired, within the constraints of adequate content coverage. The primary area of focus impacting test administration time would be related to constructed-response items, whether they be machine scored, automatically scored through artificial intelligence, or hand scored. The tradeoff with these item types is that they typically allow for the greatest depth of measurement, but deeper student engagement most often requires more time to administer these items than others.

One approach for consideration would be to start with the PARCC test blueprints as a model for ideal test length and rigor and then look for areas where compromises could be made to reduce cost and/or administration time. Pearson is intimately familiar with the ADE's past test development practices and can offer our experience developing items for the PARCC assessments and implementing custom, high-quality next generation assessments for other states to inform the ADE's design process.

Requirement

b. Identify the anticipated testing window for each assessment by grade level.

Response

ACT Aspire

Testing Windows

The ACT Aspire assessments will be administered for the first time in spring 2014. We will offer an opportunity for the summative assessments in fall 2014 while we begin rolling out the formative assessments. In subsequent years, we envision similar testing windows; however the firm dates have not been determined beyond 2014. The following figure summarizes the anticipated testing windows:

ACT Aspire Anticipated Testing Windows			
Type	**Timeline**	**Grade Levels**	**Subjects**
Summative	April 28- May 23, 2014	3 through 10	Reading, English, Writing, Math, & Science
Summative	September 15 – October 24, 2014	3 through 10	Reading, English, Writing, Math, & Science
Formative/Interim 1	Fall 2014 - TBD	3 through 10	Reading, English, Writing, Math, & Science
Formative/Interim 2	Winter 2014- TBD	3 through 10	Reading, English, Writing, Math, & Science
Formative/Interim 3	Spring 2015 – TBD	3 through 10	Reading, English, Writing, Math, & Science
Summative	Spring 2015 – TBD	3 through 10	Reading, English, Writing, Math, & Science
Summative	Fall 2015 - TBD	3 through 10	Reading, English, Writing, Math, & Science
Formative/Interim 1	Fall 2015 - TBD	3 through 10	Reading, English, Writing, Math, & Science
Formative/Interim 2	Winter 2015- TBD	3 through 10	Reading, English, Writing, Math, & Science

ACT Aspire Anticipated Testing Windows			
Type	**Timeline**	**Grade Levels**	**Subjects**
Formative/Interim 3	Winter 2016 - TBD	3 through 10	Reading, English, Writing, Math, & Science
Formative/Interim 4	Spring 2016 – TBD	3 through 10	Reading, English, Writing, Math, & Science

ACT QualityCore

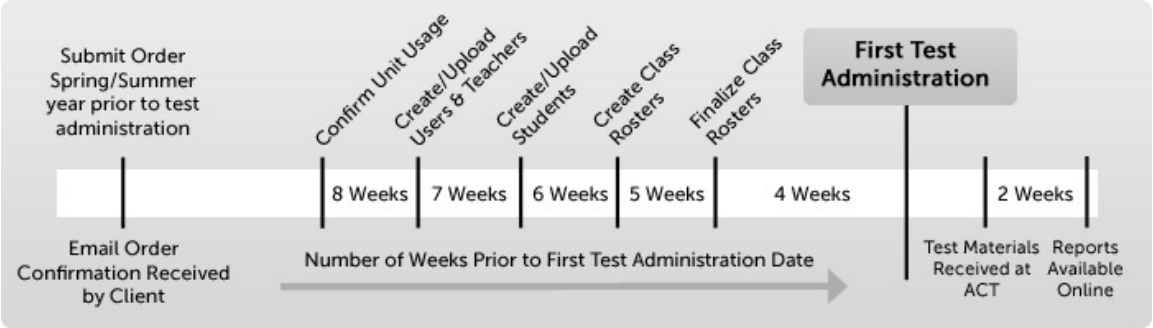
Test Calendars and System Schedule for ACT QualityCore

ACT QualityCore EOC testing has been successfully delivered on both an "off-the-shelf" and custom contract basis. ACT will work with the ADE to develop time-and-task schedules that are consistent with the operational needs of the program, both for ACT and the ADE, as well as the assessment and reporting needs of the ADE's constituent districts, schools, students, and other potential stakeholders.

Detailed Timeline

The following figure depicts a standard ACT QualityCore testing timeline for 2014-2015.

ACT QualityCore Testing Timeline



[redacted] The timeline will allow for a significant reduction in effort from the administrators of ACT QualityCore and reduction in the time needed to prepare for the exam.

The ACT

Test Dates for The ACT In-School Program

The ACT In-School program offers one fall and three spring weekday administration choices for The ACT. Each of these weekday administration choices consists of the following:

- Initial Test Date (standard time testing)
- Makeup Test Date (standard time testing)
- Two-week Accommodations Test Window (commencing on the Initial Test Date and ending on the Makeup Test Date)

ACT has established the In-School Test Date calendars shown below for the upcoming school years. Arizona would be able to choose one of the following administrations shown in the figure below. [redacted]

The ACT In-School Test Date Calendar for PBT

TEST DATE	Published Dates for PBT		
	2013-2014	2014-2015	2015-2016
Fall In-School INITIAL Test Date		Tuesday September 30	Tuesday October 6
Fall In-School MAKEUP Test Date		Tuesday October 14	Tuesday October 20
Early March In-School INITIAL Test Date	Tuesday March 4	Tuesday March 3	Tuesday March 1
Early March In-School MAKEUP Test Date	Tuesday March 18	Tuesday March 17	Tuesday March 15
Mid-March In-School INITIAL Test Date	Tuesday March 18	Tuesday March 17	Tuesday March 15
Mid-March In-School MAKEUP Test Date	Tuesday April 1	Tuesday March 31	Tuesday March 29
April In-School INITIAL Test Date	Wednesday April 23	Tuesday April 28	Tuesday April 19
April In-School MAKEUP Test Date	Wednesday May 7	Tuesday May 12	Tuesday May 3

Providing a Test Date Calendar. The ACT In-School program offers one fall and three spring weekday administration choices for administering The ACT.

Custom Assessments

Test administration windows for custom assessments could be designated to reflect many of the same considerations that drive Arizona's current February/March and April/May testing windows. Factors influencing the length of test administration windows include the number of operational forms used (to provide additional security over longer testing windows), districts' and schools' average estimated and projected capacity for online testing, total estimated testing time, and desired reporting timeframes. As we have in the past, most recently for the AZELLA program, Pearson can consult with the ADE to determine a feasible approach that best meets Arizona's needs while appropriately balancing risk against the "triple constraints" of cost, scope, and time.

Requirement

c. Describe the training needs and available training for teachers and administrators pertaining to the administration of the assessments.

Response

ACT Aspire

Customer support is paramount to successful implementation and administration of the ACT Aspire assessments. In addition to the ability of administrators and teachers to email or call our toll-free customer support team, we can provide comprehensive trainings as follows:

- Overview of ACT Aspire
- CBT system readiness (see below)
- Data management (loading organization and student data)
- Order management (ordering materials and optional services)
- Test setup and management (creating, viewing, editing, and monitoring test sessions)
- Test administration (printing authorization tickets and launching test sessions)
- Post-test interpretation and support (accessing reports and interpretation materials)

Readiness Support

For computer-based testers, customers have access to readiness support through the following tools and trainings:

- A system overview that outlines technical aspects of TestNav, the online delivery platform for ACT Aspire assessments
- An overview of the TestNav Early Warning System (EWS), which describes the integrated functionality that gives TestNav a high degree of fault tolerance and provides additional fail-safes in the event of unexpected network disruptions during online testing
- A system check tool that helps district and school technology personnel better assess the overall readiness of their local network for online testing

- Proctor caching, a Pearson-supplied software used in conjunction with TestNav to reduce bandwidth requirements for school or district networks and accelerate the delivery of test content. Proctor caching enhances the performance of existing technology in schools. No additional hardware is required

We can also provide onsite support as needed throughout this readiness phase; however, the scope and nature of such requests may incur additional charges that can be borne by or shared with the district(s) or school(s) involved.

Phone Support

Providing Responsive Customer Service to Arizona

To provide call center service to handle phone calls and emails regarding the SBCA, Pearson offers a Customer Service Center (CSC) with a strong history of support for diverse state and national assessment testing customers, including both the AIMS and AZELLA programs. Formed in 2003, our CSC services approximately 110,000 contacts per year and currently supports 58 individual assessment programs and six unique platforms/products across multiple states.

Our CSC has experienced staff members available to respond to the needs of Arizona District Test Coordinators and technical support personnel. Arizona callers receive prompt and personal service from qualified staff in sufficient numbers. Our team members are experienced and competent customer service professionals who can answer questions efficiently and thoroughly. To date in 2013, customers who completed the customer satisfaction survey indicated the CSC met or exceeded their expectations 95 percent of the time.

The Pearson CSC offers a range of services to effectively support callers from Arizona:

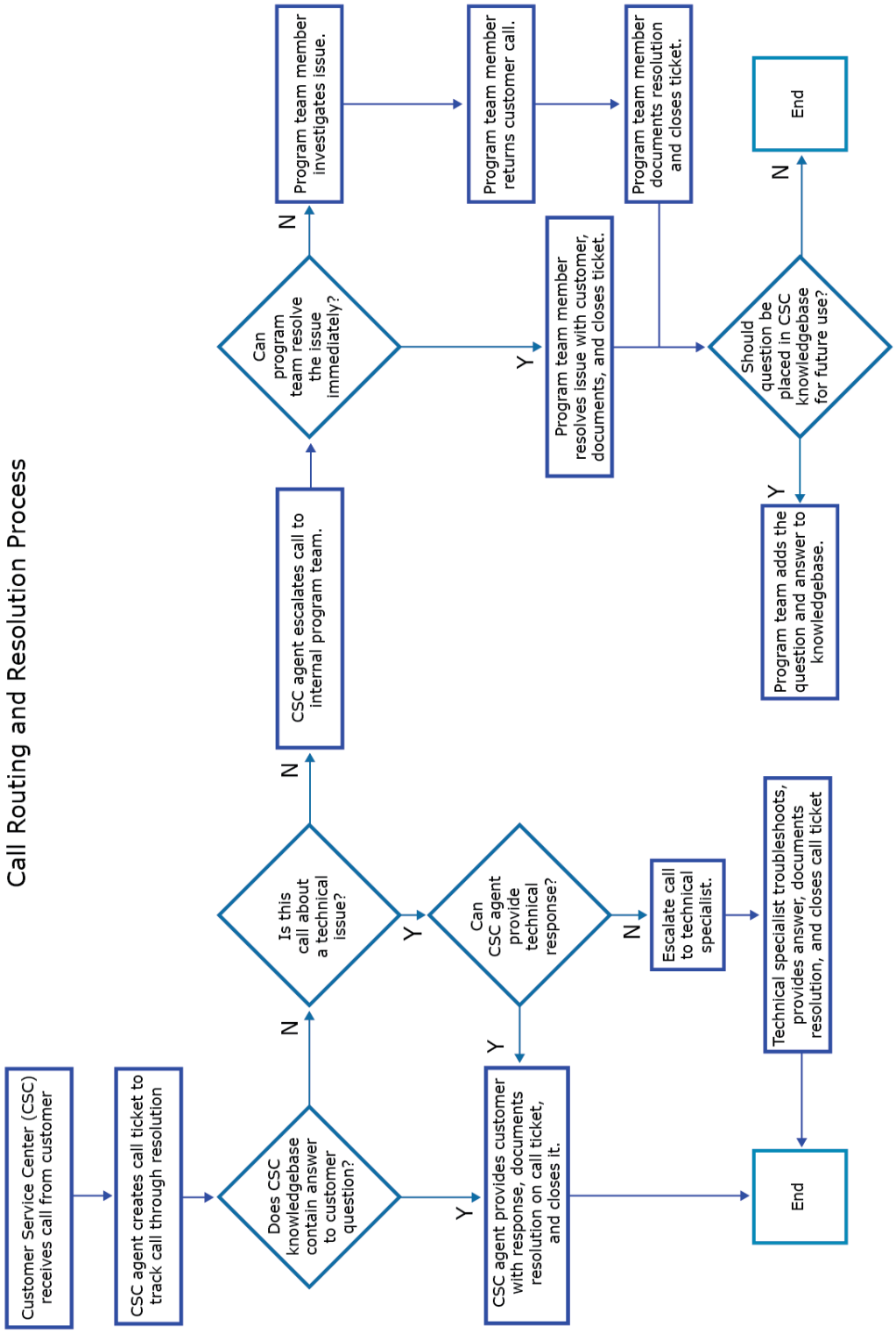
- Callers will have **access to live support** via email, chat, and toll-free phone from 7:30 a.m. to 4:30 p.m. (MST) each weekday, excluding federal holidays.
- **Specialists** with detailed program information and focused expertise in the database, network, infrastructure, and software components of Pearson's web-based services will respond to questions that cannot be resolved at the initial point of contact.
- Using **incoming call routing to manage call flow and expedite service**, we direct calls to team members best able to answer the question, including specialists to whom calls may be escalated. We als**o offer email support, i**n**cluding automated return email to con**firm receipt. For email received during off hours, we will promptly respond during the next regular business day.
- Our solution provides **access to the same knowledge base** used by CSC staff to obtain information needed to answer questions and provide information. The knowledge base is continuously updated in collaboration with the program team and available for use by callers.

Scalable Capacity for Efficient and Dependable Service

Our flexible, scalable operations allow us to provide consistent, quality service during peak periods by routing calls as needed. Operating across a network of centers in Iowa, Minnesota, and Texas, our CSC is more efficient because it is independent of a single location.

We create a contingency plan for peak periods using scheduling and forecasting information from the program team and CSC data. In periods of anticipated peak call volumes, we add trained personnel who can log in to the phone system and begin taking calls. This built-in additional capacity will reduce wait time for Arizona callers during peak volumes.

The following figure shows our call routing and resolution process.



Call Routing and Resolution. Our escalation protocol will efficiently route Arizona callers to the appropriate personnel.

Trained, Effective Staff

To meet Arizona's needs, we provide qualified and well-trained staff. CSC personnel are trained to quickly determine the scope and impact of reported problems and efficiently route and resolve them. To provide the high quality service for the SBCA, training for new CSC staff will include the following:

- Overview of knowledge, procedures and processes used for supporting our customers
- Pairing up with an experienced staff member to begin listening to customer calls and receiving instruction on phone system operation
- Detailed introduction to and instructions for use of specific websites, products and platforms
- Familiarization with tools used to house knowledge and call information
- Role playing with other staff members and practice with specific call scenarios that may be encountered
- Live customer responses with mentoring by a staff member

Staff members receive regular feedback from the supervisory staff and also receive individual call statistics, so they can see their own progress. Arizona test coordinators benefit by having access to people who can respond appropriately even when they are new to their job.

Customer Service Systems and Tools

The Pearson CSC offers the following systems and tools to effectively serve Arizona callers:

- **Telephony.** By effectively creating options from which the caller chooses, we use the Avaya telephony system (Avaya G3R Version: V15 w/CM 5.2) to route calls to CSC members trained to handle specific types of issues. This routing allows for a more efficient flow of calls to CSC staff who have the best opportunity to resolve issues during the first call.
- **Chat.** The CSC team uses LivePerson to provide real time customer support. This allows customers the flexibility to interact with agents online.
- **Knowledge Base.** The CSC team uses the Servigistics knowledge base (Kaidara Advisor 4.4) to access the information needed to answer Arizona caller questions allowing for consistent responses across team members. Because our knowledge base is continuously updated in collaboration with the program team and, as needed, approved by the ADE information that may change is promptly available to CSC staff for use in answering questions.
- **Incident Tracking.** Contacts are tracked by the HP Service Manager Client 7.11.259, a secure incident-tracking software tool tailored specifically for the Pearson CSC. With each incoming contact, CSC staff generates a ticket that enables tracking the incident through to resolution. Tickets are archived and ticket numbers are also provided to the caller. Tickets may be escalated to specialists for their documentation when handling calls to generate a full history of each issue.

- **Quality Monitoring.** We use the Verint Impact 360 Work Force Optimization Suite (v. 11) to provide high-quality service to Arizona. Through recording/monitoring of calls, CSC team members receive regular feedback from CSC supervisory staff and also receive individual call statistics, so they can see their own progress.
- **Caller Satisfaction Surveys.** To confirm we are meeting your needs, we email surveys to gain valuable information on how Arizona callers view their experience with CSC staff and our response to your concerns.

ACT QualityCore

Educator Resources

ACT QualityCore provides a wealth of fully developed educator resources to shape rigorous content and improve outcomes. Materials are based on empirical research of classrooms that have a proven success record in preparing students for college.

Providing a Research-Based Model for Rigor in the Classroom

ACT recognizes the central role teachers play in ensuring that students are prepared for future academic, workplace, and lifelong success. ACT is experienced in providing effective tools, training, and support that increases a teacher's ability to meet students' needs. The ACT QualityCore Educator Resources include tools designed to provide a comprehensive framework for rigorous coursework, the course-specific skills needed for college readiness, and strategies and models for building instructional units that are rigorous and address important skills**.**

ACT QualityCore provides educators with the following features:

- An empirically-derived identification of the essential skills leading to college readiness
- Formative and end of course assessments to measure student attainment
- Resources to improve the rigor, quality, and consistency of instruction for all students (not just the most academically able)
- Longitudinal college readiness and course rigor data for districts, schools, teachers, and students
- End of course test data that allow for consistent comparisons
- Results that can highlight the need for professional development/other support

ACT QualityCore's Educator Resources, which are aligned to the ACT Course Standards and end of course assessments, reflect a research-based approach to rigorous high school courses. These resources also have been created in collaboration with practicing, well-qualified high school teachers who are preparing their students for college and for careers at higher proportions than teachers nationally.

ACT QualityCore Educator Resources are based on empirical research of classrooms with a proven success record in preparing students for college. Materials include a program-level educator's guide, along with ACT Course Standards, a course description and syllabus, a course outline, plans for

instructional units, a model instructional unit, guidelines for developing instructional units, and an EOC test blueprint for each course.

Instructional Materials and Support

ACT's QualityCore instructional materials and training services, are outlined and described below. These materials and services are fully developed and ready for delivery.

ACT QualityCore Standard Instructional Materials

The following standard items are available for all 12 courses:

1. Educator's Guide
2. Online manuals:
    - Room Supervisor's Manual
    - Test Supervisor's Manual
    - Online User Guide
3. Online reports available at the district (where applicable), school, class, teacher, and student levels
    - Score interpretive guides for teachers.
    - Score interpretive guides for student/parent
4. End of Course Multiple-Choice + Constructed-Response MC+CR Assessment.
    - Two test booklets (one complete test form)
    - School and Group Header sheets
    - A separate student-level answer document.
    - Mathematics and Science reference sheets
    - Accommodated materials (if requested) including:
        - Braille
        - Large Type
        - Audio CD
        - Reader's Script
5. Access to course-specific Educator Resources
    - Course Description and Syllabus
    - Course Outline
    - Course Standards
    - Rigor and Relevance Template
    - Instructional Unit Plan
    - Guidelines for Developing an Instructional Unit

- One Sample Instructional Unit
- End of Course Test Blueprint
- Formative Item Pool: Linked to specific Course Standards
- Average of 300 multiple-choice and constructed-response items
- 4–6 Benchmark Assessments
- College Readiness Standard—Course Posters

The ACT

Training Testing Staff

A primary goal with any ACT In-School program weekday administration is a carefully planned implementation strategy that includes an initial orientation (Implementation Training) for school and district officials. This is followed by detailed training (Test Administration Training) of district and school testing staff (Test Supervisors, Back-up Test Supervisors, and Test Accommodations Coordinators) as well as a specific training focusing on the test accommodation application process. ACT also provides webcast modules tailored to the Room Supervisors and Proctors providing the level of detail needed to successfully administer The ACT within his or her assigned testing room.

Following is a detailed description of ACT's standard training model for the ACT In-School Testing program.

Implementation Training

Implementation Training is conducted once in the first year of the contract with refresher training in future years. The focus of Implementation Training is the orientation of school and district staff to ACT In-School program. Implementation training is conducted by ACT staff and includes the following:

- Overview of In-School program testing
- A review of the calendar, decisions points, key events and due dates
- Expectations, roles and responsibilities

ACT will work with the ADE to provide implementation training at regular and appropriate intervals throughout the contract period.

Test Accommodations Training

ACT will make available via an online webcast a Testing Accommodations Coordinators (TACs) Training video. This Webcast covers in detail the accommodations process, documentation, and procedures for requesting ACT-Approved Accommodations. It also covers the procedures TACs must follow to order State-Allowed accommodations materials for students who do not meet the ACT-approved test accommodations criteria (if the state chooses this option).

In addition to the webcast, ACT will provide a question and answer session for TACs during the fall of each cycle and in advance of the deadline for ACT-Approved Accommodations Applications.

Test Administration Training

Test Administration Training will provide detailed instructions for all steps of the test administration, including initial planning, identification and training of room supervisors and proctors, receipt and check-in of secure materials, secure storage requirements, test day arrangements, online administration procedures, documentation of required procedures, recognition of irregularities, accounting for and return of materials, and plans for makeup testing. ACT staff will provide answers to frequently asked questions and help appointed Test Supervisors, Back-up Test Supervisors, and Test Accommodations Coordinators assess their readiness for the administration.

To offer maximum flexibility to state educators, Test Administration Training is delivered through an online pre-recorded webcast followed by two live question and answer sessions in advance of the test dates and in time for testing staff to successfully complete pretest responsibilities. It is expected that appointed Test Supervisors, Backup Test Supervisors, and TACs will participate in the training.

To achieve acceptance of In-School Testing results as comparable to national test date scores, ACT has assured colleges that training and directions for paper and pencil administrations are coming directly from ACT. In addition, ACT has provided assurances to the NCAA that scores achieved through paper and pencil In-School Testing are comparable for use in determining eligibility for Divisions I and II NCAA Initial-Eligibility. [redacted] As such, ACT administered training helps to achieve the following:

- Provide consistent training across schools
- Support the “high stakes” nature of The ACT and “college reportable” scores
- Provide the appropriate/consistent response to trainee questions
- Eliminate misinformation in training that may potentially lead to a misadministration
- Provide scores that colleges and the NCAA have been assured are comparable to national ACT scores.

Training of Room Supervisors and Proctors

Based on the number of students testing at a school, Test Supervisors and TACs will need to identify and train additional school staff to support the test day administration and accommodations testing. Each testing room will require a Room Supervisor and depending on the number of students in each room may require additional proctors during the course of testing. It is required that prior to test day Test Supervisors and TACs conduct (school-level) test administration training for all personnel involved in testing (e.g., Room Supervisors and Proctors.) ACT provides an online webcast that focuses on roles and responsibilities of Room Supervisors and Proctors.

Custom Assessments

Training

Training support for custom assessments provided by Pearson would follow a scope and sequence already familiar to most Arizona District Test Coordinators and Technology Coordinators. For example, Arizona schools and districts will soon be participating in full-scale training for the upcoming PARCC online field test administration through PearsonAccess and TestNav. The scope of this training will encompass all phases of online testing readiness, including both technology-related and administrative duties. Although much will be familiar to District Test Coordinators due to similarities to Arizona's current use of PearsonAccess for AIMS and AZELLA, they will also be experiencing new functionality for the first time, and an entirely separate track will be delivered for local Technology Coordinators.

Initial training sessions are being conducted in live, large group settings in three locations across the state, with additional online follow-up sessions to be conducted more frequently via WebEx.

To provide further training in modes that best suit different learning styles, Pearson maintains a library of online modules pertaining to specific test administration and delivery topics that can be reviewed at any time, at each individual's own pace, and can be reused by District Test Coordinators and Technology Coordinators in train-the-trainer mode with other local staff according to each district or school's own schedule.

These delivery methods are similar to those we have provided in conjunction with the ADE in the past and have proven to be effective both in terms of time and cost efficiency. A detailed training plan for implementing custom SBCA assessments for Arizona would be developed similar to those used in the past and the one described above, enabling the ADE to leverage much of the initial training that will have already taken place. This would allow districts and schools to maintain the momentum established through this coming spring's field testing, reducing the overall training effort required to successfully launch Arizona's new assessment system.

5. Assessment Standards Setting and Scoring

Requirement

a. Describe the standards setting process.

Response

ACT Aspire and The ACT

Standard Setting

ACT encourages an evidence-based approach to standard setting that uses established relationships between test scores and college and career readiness, among other evidence. This approach can supplement or replace traditional standard setting methods that are based on panelist judgment of student performance data on actual test items. With this approach, proficiency levels are defined (or informed) by college and career readiness indicators. Because ACT's college and career readiness indicators are defined for each grade level and reflect typical student growth, this approach promotes continuity across grade levels in proficiency levels. It also provides that Arizona's proficiency levels are set at nationally and internationally competitive levels.

Justifying the Standards Verification Approach

To provide the ADE a standard setting or standards verification approach with a clear justification of the particular method, we would conduct a cut point review and evaluation process with Arizona educators. The process would examine how well the initial (established) achievement level descriptors reflect the content of Arizona's College and Career Ready Standards and how consistent the cut points are with Arizona grade-level expectations (GLEs).

The process would use actual performance data on a representative set of assessment items mapped to Arizona's standards. The item performance data would allow participants to link assessment scores to their own judgments of Arizona's designations of below basic, basic, proficient, and advanced.

Differences between the established cut points and those matching Arizona's achievement level designations would be reconciled with a crosswalk. Additional work may be needed to confirm that revised cut points result in a steady progression across grade levels.

Training, staff, and psychometric support needed for the cut point validation and logistical requirements would be provided, as well as a technical report that documents the entire process and results.

A Sound Basis for Established Achievement Levels

The ACT College Readiness Benchmarks (in English, mathematics, reading, and science) are the scores needed to have at least a 50 percent chance of earning a "B" or higher grade in aligned first year credit-bearing college courses (English composition, college algebra, social science courses, and biology). The ACT College Readiness Benchmarks were updated in 2013 and are based on data from 214 colleges and over 230,000 students from across the United States.

ACT Aspire's Readiness Benchmarks for earlier grade levels are defined as the scores needed to be on target to meet the ACT Benchmarks by spring of grade 11. ACT Aspire's Readiness Benchmarks will be available in spring 2014, allowing users to monitor student growth relative to peers and relative to the college and career readiness targets. ACT Aspire will include a two-year "predicted path" that will project student performance based on typical student growth between summative assessments.

Other existing resources that can be used in the standard setting process include ACT's College Readiness Standards, alignment to National Assessment of Educational Progress (NAEP) and other current achievement levels in use, test score percent-at-or-below tables (normative information), and relationships of ACT score points to probability of college course success and college degree attainment. These resources may be especially useful in determining additional performance levels above and below the proficient level. ACT also offers to work with Arizona on an analysis of ACT scores and success in first-year courses at Arizona's postsecondary institutions (this study would require Arizona's colleges and universities to provide existing first-year course grade data for ACT-tested students).

Options for Traditional Arizona-Specific Standard Setting

Should Arizona request traditional standard setting methods that are based on panelist judgment of student performance on actual test items, Pearson and ACT are well-equipped to lead. In addition to numerous other statewide assessments, Pearson has conducted four standard settings for Arizona assessments within the last three years alone. ACT also has a great deal of experience in standard setting and has been responsible for many of the advances in this area. ACT was the sole contractor used by the National Assessment Governing Board (NAGB) for setting standards on the NAEP between 1992 and 2009. The table below lists some of ACT's experience in standard setting.

Client	Method	Course/Grade	Client	Method	Course/Grade
NAGB	Modified Angoff	Civics/4–8-12	NAGB	Modified Angoff	History/4–8-12
NAGB	Modified Angoff	Writing/4–8-12	NAGB	Modified Angoff	Geography/4–8-12
NAGB	Modified Angoff	Math/4–8	NAGB	Modified Angoff	Science/4–8-12
NAGB	Mapmark	Math/12	NAGB	Mapmark	Economics/12
Illinois	Modified Angoff	Math/11	Illinois	Modified Angoff	Science/11
Illinois	Mapmark	Writing/11	Illinois	Modified Angoff	Social Science/11
DoDEA	Mapmark	Biology	NAGB	Mapmark	Science/4-8-12

ACT and Pearson would work with the ADE to establish the different achievement levels, with possible modifications of the process to incorporate the existing college readiness benchmarks and the additional resources described above. The standard setting method we would recommend, Mapmark, is a modification of the Bookmark procedure (further details on the Mapmark procedure are available upon request from ACT). Panelists would be recruited by the state, and may include, at the state's discretion, state representatives, content-area assessment specialists, LEA assessment directors, principals, teachers, and content-area specialists. Facilitators would be Pearson and ACT content area specialists, statisticians, and psychometricians. In addition to the panelists, each content area room would have a technical facilitator and a content facilitator. Two sessions would be conducted sequentially over two days, with 1–2 content areas addressed in each session. Data would be collected at a subsequent meeting and with preliminary results of the standard setting available for internal review at the conclusion of that meeting.

ACT QualityCore

Standard Setting

Because ACT QualityCore is an existing assessment with existing predictors to college and career readiness indicators, we suggest conducting a cut point validation rather than a full achievement level setting, which will lower the overall cost of the solution while still assuring predictive validity of the assessment.

If Arizona should desire a separate standard setting for its use of the ACT QualityCore assessments, ACT and Pearson can easily accommodate this request. Similar to the approach we described for ACT Aspire and The ACT, we would suggest an evidence-based approach for standard setting for ACT QualityCore exams. Existing standard setting resources for ACT QualityCore can supplement or replace a traditional standard setting effort.

Student performance benchmarks for ACT QualityCore assessments have been empirically derived. Because estimated ACT or Plan score ranges are provided on each ACT QualityCore student overview report, and because college readiness benchmarks already exist for ACT and Plan scores, the student overview report includes student performance benchmarks that predict success in first-year credit-bearing postsecondary courses. The estimated Plan/ACT score range is a prediction of how a student is likely to score on the Plan or The ACT if Plan or The ACT is taken at the same time as the ACT QualityCore assessment. The Plan/ACT score intervals for each ACT QualityCore scale score are provided in the ACT QualityCore Technical Manual. Defining proficiency based on the ACT QualityCore performance benchmarks allows continuity in the definition of proficiency—both between different ACT QualityCore exams and between ACT QualityCore, ACT Aspire, and the ACT College Readiness Assessment. Proficiency definitions would be linked to college and career readiness targets.

Other existing resources that can be used in the standard setting process include alignment to other current achievement levels in use, test score percent-at-or-below tables (normative information), relationships of ACT QualityCore score points to probability of meeting the ACT College Readiness Benchmark in the same subject area, and relationships of ACT QualityCore score points to the probability of earning “A” grades in the high school course.

Should the ADE wish to explore additional standard setting for the ACT QualityCore assessments, ACT is well-qualified to assist in that process, as described above for ACT Aspire and the ACT College Readiness Assessment. For ACT QualityCore, we would also suggest the Mapmark method.

Custom Assessments

Standard Setting

As previously mentioned, if Arizona were to use sufficient numbers of items licensed from PARCC (or SBAC), it would likely be possible to create custom operational forms that could be equated to either consortium's operational assessments. In this case, if Arizona were willing to accept these assessments' corresponding performance standards, an Arizona-specific standard setting may not be necessary, particularly if there was a high degree of similarity to Arizona's custom blueprints.

If Arizona should desire to set its own performance standards, Pearson psychometric staff have significant experience in designing and implementing standard setting studies for a number of large-scale assessment programs.

In recent years, we have conducted standard setting studies employing a variety of approaches, including Item-Mapping (aka Bookmark), Body of Work, Modified Angoff, Reasoned Judgment, and Contrasting Groups. We have combined multiple approaches, developed customized procedures, and modified materials and panel training as needed to tailor a given approach to the needs of our state clients. For a custom SBCA, Pearson would consult with the ADE and its Technical Advisory Committee (TAC) on the advantages and disadvantages of various standard setting approaches, to devise a final methodology based on the desired test blueprint, and in light of the multiple high-stakes uses associated with Arizona's future assessments.

The following figure shows some of the standard setting activities Pearson has conducted in recent years, focusing primarily on the Item Mapping (aka Bookmark) method Arizona has used most frequently for the AIMS and AZELLA. We have also included a small sample of standard setting meetings we have conducted using other methodologies to demonstrate our experience and flexibility in meeting our clients' standard-setting needs.

Pearson Standard Setting Experience					
Program Name	**Contract Dates**	**Test Type**	**Subjects**	**Grades**	**Standard Setting**
Arizona AIMS	2009–2014	EOG (CRT)	Mathematics	3–8, 10	Item Mapping
Arizona AIMS	2009–2014	EOG (CRT	Writing	5,–7, 10	Item Mapping
Arizona AZELLA	2011–2014	Placement	ELL	Kindergarten	Modified Angoff
Arizona AZELLA	2011–2014	Placement, Reassessment	ELL	K–12	Item Mapping

Pearson Standard Setting Experience					
Program Name	**Contract Dates**	**Test Type**	**Subjects**	**Grades**	**Standard Setting**
Florida Comprehensive Assessment Test (FCAT)	2000–2005	EOG (CRT)	Reading, math science writing	3–Adult 5, 8,11 4, 8, 10	Item Mapping (2005) Body of Work (2006) 5, 8 science
Georgia End of Course Test (EOCT)	2001–2011	EOC (CRT)	Literature (9th grade & American), algebra, geometry, physical science, biology, US history, economics, math I, math II	9–12	Modified Angoff (2005) English, science
Hawaii State Alternate Assessment	2005–2007, 2009	ALT (Portfolio)	Reading, math, science, writing	3–8, 10–11	Modified Bookmark (2006) math, reading
Maryland School Assessment Science (MSA Science)	2006–2011	EOG (CRT)	Science	5, 8	Item Mapping (2007)
Michigan Educational Assessment Program (MEAP)	2005–2008	EOG (CRT)	ELA, math, writing science social studies	3–8 5, 8 6, 9	Item Mapping & Body of Work math, science, reading/ELA, social studies K–12 grade levels per content area (2006)
Michigan English Language Proficiency Assessment (ELPA)	2005–2008/10	ELP (CRT)	Listening, reading, writing, speaking	K–12	Item Mapping (2006)
Minnesota Comprehensive Assessments (MCA)	1998–2011	EOG (CRT)	Reading math science writing	3–8, 10/11 5, 8, HS	Modified Angoff writing, reading, math, 3, 5, 10, 11 1998, 2004 Item Mapping Science (2008)
Minnesota Test of Academic Skills (MTAS)	2007–2011	ALT (portfolios)	Reading math science	3–8, 10 3–8, 11 5, 8, HS	Item Mapping 2007, 2008
Mississippi MCT2/MSATP	2007-2008	ALT (Observation of assessment tasks)	Language Arts, math algebra I, English II	3-8 EOC (HS)	Item Descriptor (ID) Matching (2008)
New Mexico Standards Based-Assessment (NMSBA)	2003–2009	EOG/HS (CRT/NRT)	Reading, writing, math, science, social studies	3–8, 11	Modified Angoff (2005-2009) Item Mapping for Writing (2008–2009)

Pearson Standard Setting Experience					
Program Name	**Contract Dates**	**Test Type**	**Subjects**	**Grades**	**Standard Setting**
New Mexico English Language Proficiency Assessment (NMELPA)	2006–2009	ELP (NRT-augmented)	Pre-reading, pre-writing, reading, writing, listening, speaking	K–12	Modified Angoff (2006) Item Mapping (2008)
New York Assistance in Educational Data Analysis and Evaluation	2002–2005	EOC (CRT)	Math, algebra, geometry, English, US history, government, global history, geography, physics, chemistry, biology, earth science	9–12	Item Mapping (2003) physics
North Carolina End of Course Tests	2006–2007	EOC (CRT)	Algebra, algebra II, geometry	HS	Item Mapping (2006) online and alternate assessment computer skills
Oklahoma End-of-Instruction Program (EOI)	2007–2008 Renews	EOC (CRT)	English 2 & 3, algebra 1 & 2, geometry, biology I, US history	9–12	Item mapping (2008) algebra II, geometry, English II & III, biology I, US history
Oklahoma School Testing Program	2001–2005	EOG (CRT)	Reading, math, science, US history, writing geography	3, 5, 8 7 (on-line)	Item Mapping (2003, 2004, 2005)
South Dakota (STEP-A)	2005–2014	ALT (rating scale and supporting evidence)	Math, reading science	3-8, 11 5, 8, 11	Modified Hofstee (2006)
Texas Assessment of Knowledge and Skills (TAKS)	2005–2010	EOG/EXIT (CRT)	Reading, writing, math, social studies, science, ELA Spanish versions: Reading, math, writing, science	3–9 4, 7 3–11 8,10–11 5, 8, 10–11 10–11 3–6 4 5	Item Mapping (2002, 2008)
Texas Assessment of Knowledge and Skills -Alternate (TAKS-Alt)	2005–2010	ALT (teacher observation of assessment tasks) (CRT)	Reading, writing math, social studies, science, ELA	3–9, 4, 7 3–10/11 8,10–11 5, 8, 10–11 10–11	Body of Work (2008) Extended Angoff / Performance Profile (2009)
Utah Criterion Referenced Tests	2004–2009	EOG/HS (CRT)	ELA	2–11	Item Mapping ELA 3, 7, 10 (2004)

Pearson Standard Setting Experience					
Program Name	**Contract Dates**	**Test Type**	**Subjects**	**Grades**	**Standard Setting**
Virginia Standards of Learning Program (VASOL) End-of-Grade Assessment	2000–2010	EOG (CRT)	Math, reading, science, history, writing	3–8	Angoff, Item Mapping (2001)
Virginia Standards of Learning Program (VASOL) End-of-Course Assessment	2001-2014	EOC (CRT)	Algebra I & II, biology, chemistry, earth science, English, reading, geometry, Virginia and US history, world geography, world history I & II	Middle school, 11	Angoff, Item Mapping (2002)
Virginia Grade Level Alternate Program VGLA)	2006–2010	ALT (Portfolio)	Writing, reading, math, science, history	3–8	Reasoned Judgment (2006)
Virginia Substitute Evaluation Program (VSEP)	2005–2010	ALT (portfolio)	Math, reading, science, history, writing	3–8 3, 5, 8 3	Item Mapping (2007) social science, history 3–8
Washington Language Proficiency Test (WLPT-II)	2005–2010	ELP (Augmented NRT: SELP)	Reading, writing, listening, speaking	K–12	Item Mapping (2006)
Proficiency Assessments of Wyoming Students (PAWS)	2004–20009	EOG (CRT)	Math, reading, science, writing	3–8, 11	Item Mapping (2005, 2008)
Wyoming English Language Learners Assessment (WELLA)	2004–2008	ELP (Augmented NRT)	ELL	K–12.	Item Mapping (2005)
American Diploma Project (ADP) Algebra I End-of-Course Exam	2007–2011	EOC (CRT)	Algebra I	10/11	Yes / No (Modified) (2009), using panelists selected from higher education

Pearson Standard Setting Experience. Arizona can rely on our psychometric staff and their significant experience with standard setting studies for a number of large-scale assessments.

Requirement

b. If already established, describe the performance levels and the performance level descriptors for each category.

Response

ACT Aspire, The ACT, and ACT QualityCore

Performance Level Descriptors

As described above, ACT's College Readiness Benchmarks (for the ACT), ACT Readiness Benchmarks (for ACT Aspire), and ACT QualityCore performance benchmarks provide established performance level cutoffs that could be used as indicators for proficiency. Other existing resources (e.g., normative information; test score relationships with success in college courses, high school course grades; and chances of meeting ACT College Readiness Benchmarks) may be useful for setting other performance levels.

ACT's College Readiness Standards provide descriptors of the knowledge, skills, and abilities of students at different levels for English, mathematics, reading, and science. Using the common score scale used by the ACT College Readiness Assessment, ACT Explore, and ACT Plan, the College Readiness Standards are reported for six score levels.

Because concordance tables will be developed for linking ACT Aspire assessments for grades 8-10 to ACT Explore and ACT Plan, the existing College Readiness Standards provide performance level descriptors for ACT Aspire for grades 8-10. Because ACT QualityCore scores are linked to ACT or Plan scores, the College Readiness Standards also provide performance level descriptors for ACT QualityCore scores, indirectly. ACT is currently planning the development of the extension of the College Readiness Standards down to grade 3 for English, reading, mathematics, and science.

Should the ADE determine that the ACT-provided performance level descriptors need to be supplemented, ACT and Pearson are well-equipped to assist in that process. Development of performance level descriptors can be included in the Mapmark standard setting process following the administration of the assessments to a sample of students.

Custom Assessments

Performance Level Descriptors

Similar to decisions around standard setting, Arizona could choose to adopt (and possibly modify) college and career ready performance level descriptors used by either the PARCC or SBAC consortia, either as the basis for equated scores or as the foundation for an Arizona-specific standard setting.

If desired, the ADE could also create Arizona-specific performance level descriptors for the SBCA, relying on its established processes for creating performance level descriptors for the AIMS and AZELLA programs, a process with which Pearson is already familiar. In this instance, researchers from Pearson's

Center for College and Career Readiness, part of the Pearson Research and Innovation Network, could also assist in identifying or developing appropriate new benchmarks for measuring college and career readiness according to Arizona preferences.

Requirement

c. Describe the score reports available to teachers, students and parents. The description should include:

   i. How the reports illustrate a student's progress on the continuum toward college and career readiness, grade by grade, and course by course; and

   ii. How the reports are instructionally valuable, easy to understand by all audiences, and are delivered in time to provide useful, actionable data to students, parents, and teachers.

Response

ACT Aspire

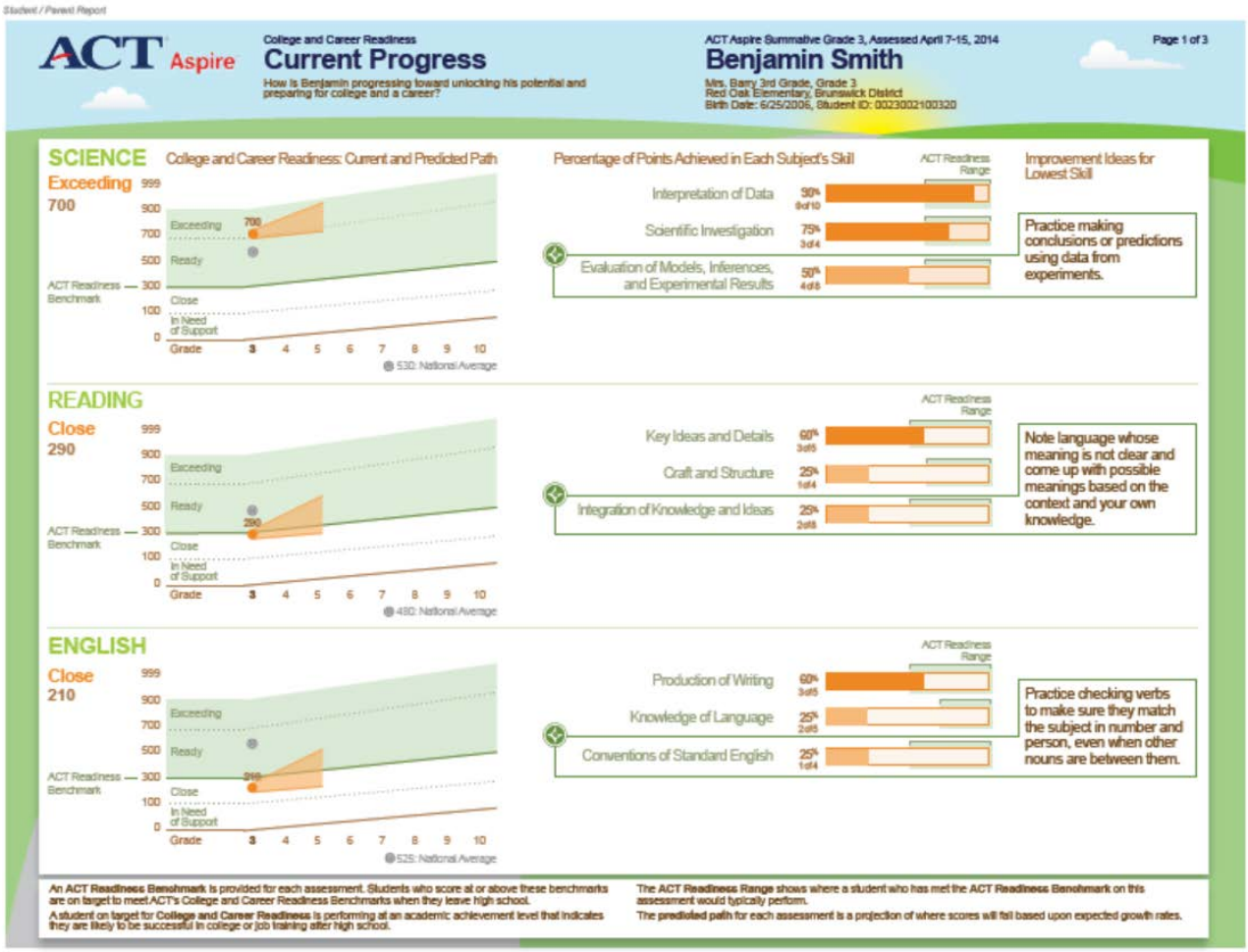
Identifying a Student's Predicted Path

Parents and educators in Arizona are interested in knowing how students are proceeding toward college and career readiness. The ACT Aspire reporting will not only report growth between summative assessments, but also identify a student's predicted path and the gaps that might exist between that path and college readiness. Reporting will focus on the progress toward college and career readiness rather than a single result, highlighting the paths to improvement.

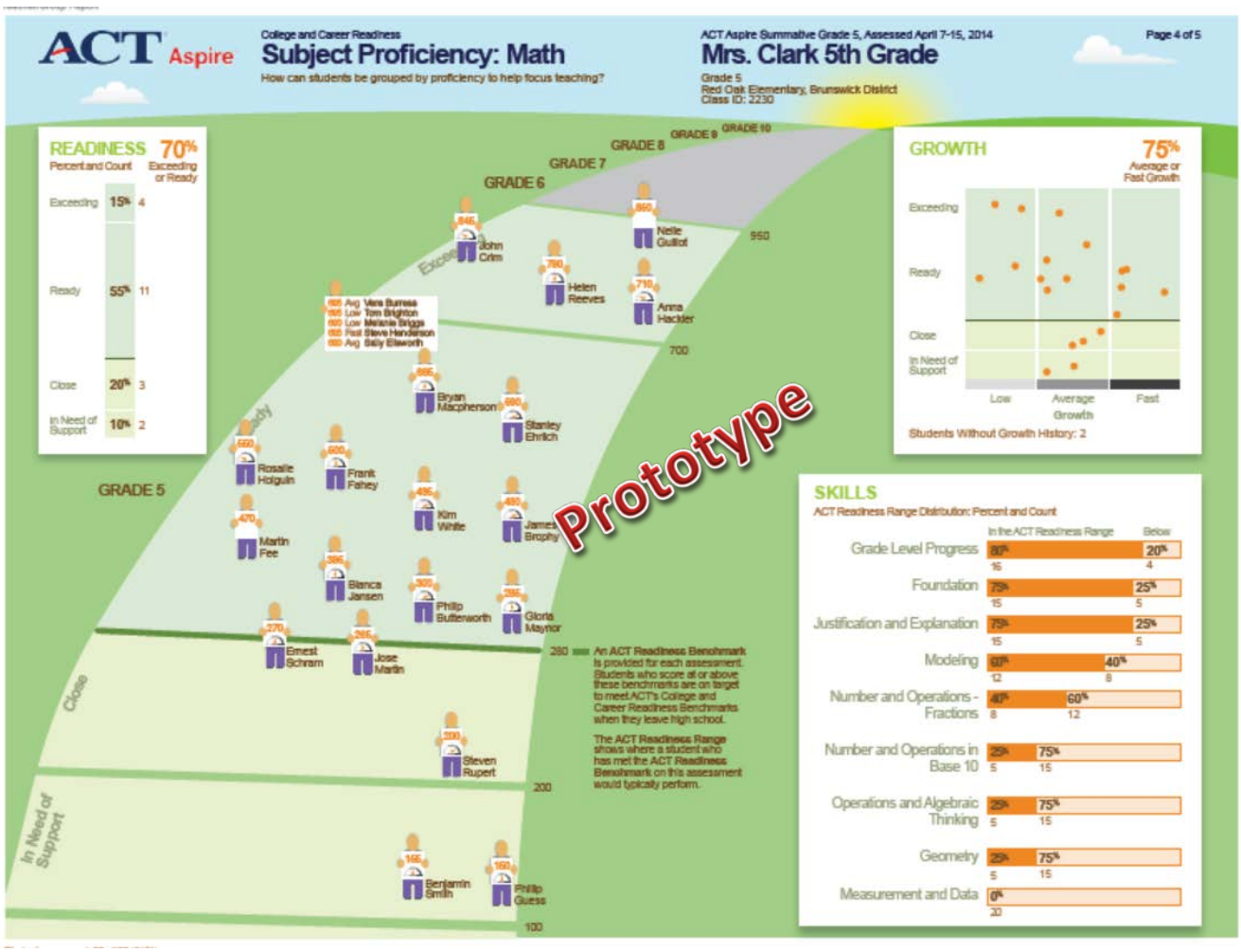
Additionally, reporting scores and scales include the following benefits:

- English, math, reading, and science reported using a three-digit vertical scale for grades 3–10, which is linked to the ACT test scale through the inclusion of ACT Readiness Benchmarks at each grade level
- New readiness measures within and across subjects to include ELA scores for those who test in English, reading, and writing; reading text complexity scores; and work readiness scores for those who test in reading and math
- A STEM score available for schools in Arizona that elect to test with both math and science
- Standards-based reporting that includes scores in reporting categories based on the ACT College Readiness Standards and aligned to the Arizona College and Career Ready Standards
- Progress charts showing student longitudinal growth versus the ACT Readiness Benchmark for each subject
- National norms
- ACT Aspire composite scores and predicted ACT subject/composite scores (grades 9 and 10 only)
- Classroom and aggregate reports with aggregated presentations of same student score report data in easy-to-understand visuals

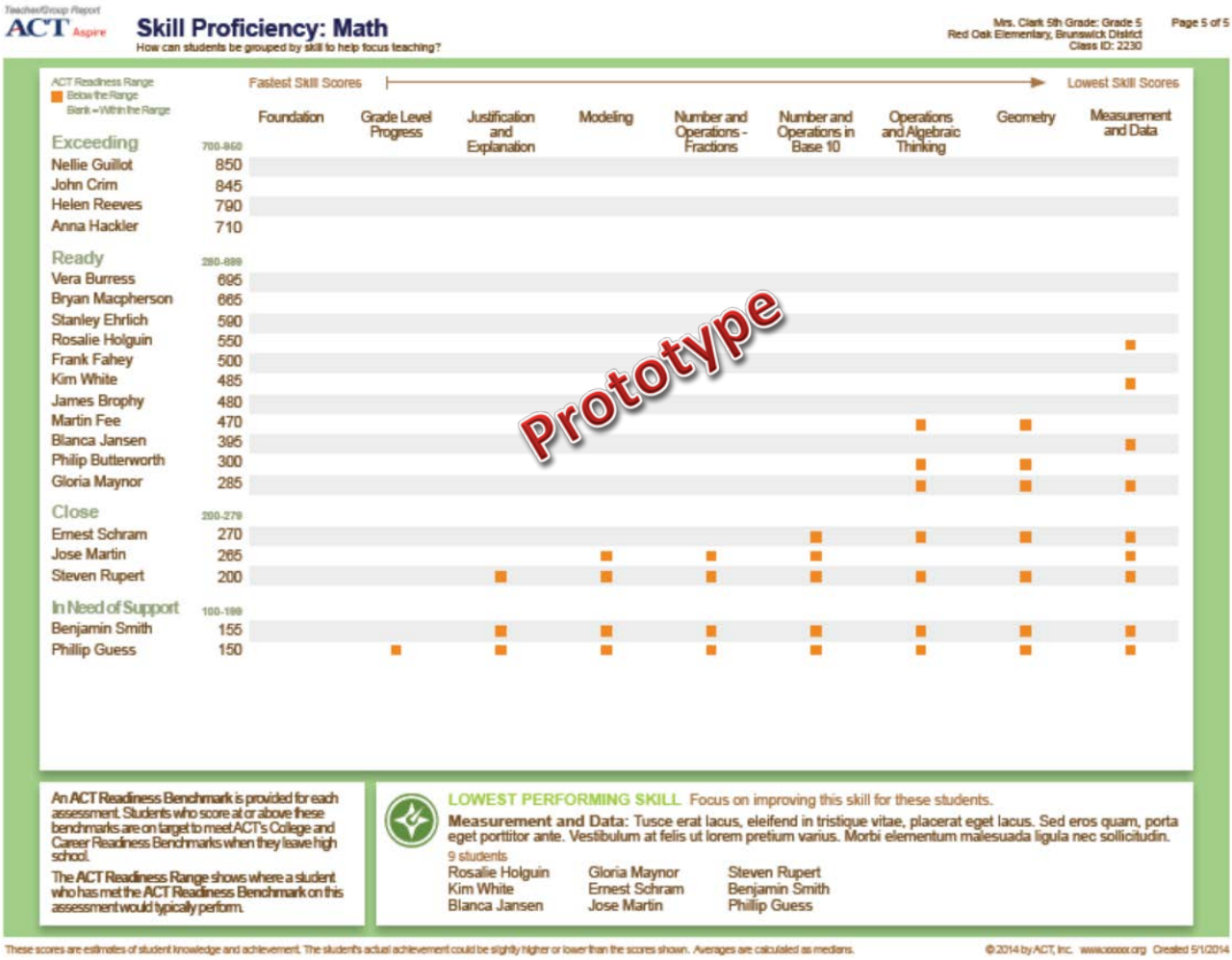
Prototype Reports for ACT Aspire



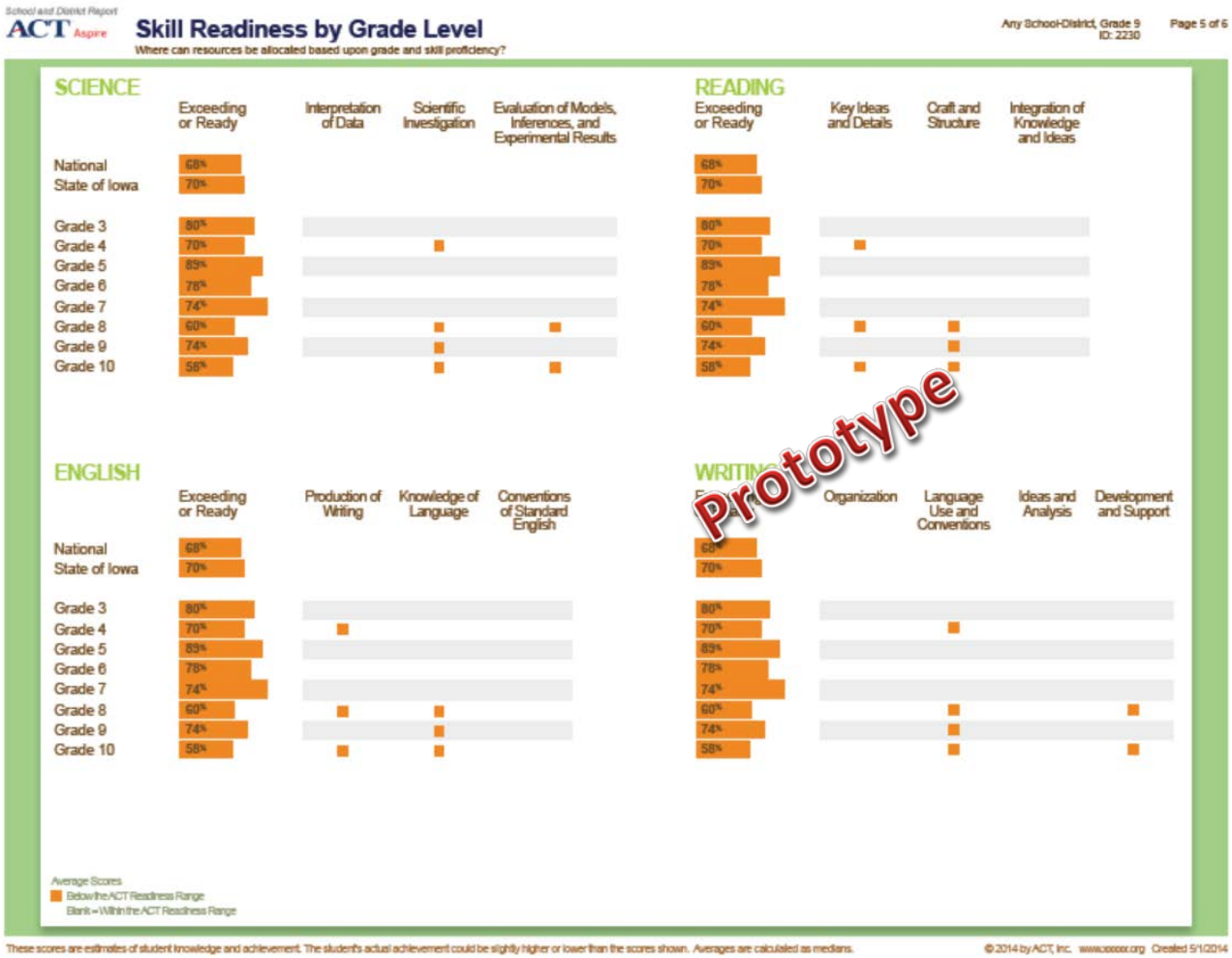
Providing the Individual Student/Parent Report. This prototype shows the main examinee report for ACT Aspire at launch, designed primarily to facilitate conversation between educator, parent, and student concerning the student's academic achievement.



Enabling Understanding of Overall Classroom Performance. The Educator/Classroom reports will enable an educator to understand the class's overall performance and growth and quickly identify the areas where a group of students are meeting (or not meeting) the benchmarks.



Demonstrating Skill Proficiency. This prototype is designed to assist educators in quickly identifying students with skill gaps, as well as identifying opportunities for curriculum adjustment or professional development.



Showing Skill Readiness by Grade Level. This prototype of school/district/state reporting provides an at-a-glance identification of reporting categories/skill areas of opportunity by grade across the cohort. It is designed to assist educators in quickly identifying grade level opportunities for improving skill gaps, as well as identifying opportunities for curriculum adjustment or professional development.

Guide to Interpreting Results

Arizona parents, educators, schools, districts, and postsecondary institutions will receive reporting services that include high quality and dependable support. These services include detailed data interpretation information.

Our reporting philosophy follows these guiding principles, based on User-Centered Design:

- The dynamic online user experience will provide the primary platform for reporting for ACT Aspire results.
- Paper reports also will be available at all levels (student, classroom, aggregate); however, deeper and richer data and insights exist in the online environment.
- Reporting is designed to be self-service, to enable customers to access their data at the time and place of their choosing after completion of scoring.

- The ADE and Arizona districts and schools will have access to downloadable PDFs of pre-packaged standard reports, including student score reports and aggregate reporting for schools, districts, and states.
- Aggregate data files of student results will be available in secure electronic format for download to Arizona student information systems.
- ACT Aspire reporting will be accessible via computer for the initial product launch and possibly via tablet in the future.

These guiding principles are supported by making available to online users detailed drill-down information that aids in measurement interpretation. The goal is to enable Arizona users to easily access the information they need to interpret assessment results.

ACT Aspire will provide customer focused materials to help parents, teachers, and students better understand what is needed for continual academic growth of students. The focus will be on the student's journey and on the progression towards readiness. Such support materials include an ACT Aspire Guide for Student and School Reports, and Understanding Your ACT Aspire Results.

ACT QualityCore

ACT QualityCore assessment score reports are delivered online through a secure, user-friendly interface that represents individual and group achievement data relevant to the user's role.

- For teachers—Student , Class, and Teacher Reports
- For district administrators—Student, Class, Teacher, School, and District Reports
- For state administrators—Student, Class, Teacher, School, District, and State Reports

Reports provide local, state, and national comparisons of students' performance within each course, as well as evaluate students' progress toward college readiness on a course-by-course basis.
ACT continues to provide a consistent scale (number) that accurately communicates student performance. In order to maintain a consistent metric year after year, ACT provides a standard ACT QualityCore scale score ranging between 125 to 175 for EOC exams.

In addition, each student's individual ACT QualityCore subject score includes a projected score to either PLAN or The ACT, with a determination of whether the student's ACT QualityCore subject score puts them on the path to achieving the corresponding ACT College Readiness Benchmark Score.
Student, school, and district reports will be available on the ACT QualityCore site 10 business days after receipt of valid answer documents. Student reports may be printed and delivered in hard copy format.

ACT QualityCore Reports and Access			
Name of Report	**District Access**	**School Access**	**Teacher Access**
Roster Report–Final by Subscore	Provides QC scores and subscores at the district, school, class, and student levels. User has ability to drill down to school, class, and teacher reports.	Provides QC scores and subscores at the school, class, and student levels. User has ability to drill down to class, and teacher reports.	Provides class QC score, student QC scores, and QC subscores. Teacher's access is limited to assigned class rosters
Roster Report–Final by Gender	Provides average QC score for total population and average QC score by gender at the district, school, class, and student levels. User has ability to drill down to school, class, and teacher reports.	Provides average QC score for total population and average QC score by gender at the school, class, and student levels. User has ability to drill down to class, and teacher reports.	Limited to class average QC score and average QC score by gender at the class level. Teacher's access is limited to assigned class rosters
Roster Report–Final by Ethnicity	Provides average QC score for total population and average QC score by ethnicity at the district, school, and class, and student levels. User has ability to drill down to school, class, and teacher reports.	Provides average QC score for total population and average QC score by ethnicity at the school, class, and student levels. User has ability to drill down to class, and teacher reports.	Provides average QC score for total population and average QC score by ethnicity at the class level. Teacher's access is limited to assigned class rosters
Overview Final Score Report	Provides final QC score (session 1 + session 2) at the district, school, class levels. User has ability to drill down to school, class, and teacher reports.	Provides final QC score (session 1 + session 2) at the school and class levels. User has ability to drill down to class and teacher reports.	Provides final QC score (session 1 + session 2) at the class and student levels. Teacher's access is limited to assigned class rosters
Student Report	Individual student score report containing final QC score, percent of students at the school/district/state levels performing at or below the student's final score, college readiness report, and subscores. District's access is limited to students enrolled in the district.	Individual student score report containing final QC score, percent of students at the school/district/state levels performing at or below the student's final score, college readiness report, and subscores. School's access is limited to students enrolled in the school	Individual student score report containing final QC score, percent of students at school/district/state levels performing at or below the student's final score, college readiness report, and subscores. Teacher's access is limited to assigned class rosters

The ACT

ACT test results will be reported in a variety of ways:

1. The ACT Standard College Entrance Reporting:

Currently, ACT reports for students who earn college reportable scores will be distributed to individual students, to the high school for which a valid reportable high school code has been provided, and to the colleges or universities requested by the student as indicated in the following chart. The current standard college entrance reports additionally incorporate important non-cognitive data elements that will be collected as part of the In-School testing process. For instance, self-reported course-plans, educational plans, and student responses to career questions that will help students and schools with postsecondary education and career planning options.

Note: Standard College Entrance Reports are delivered as processed and are not subject to the reporting schedule for Standard In- School reporting.

ACT will make the following standard reports available for those students tested through In-School Testing and who received college-reportable scores:

Standard ACT College Entrance Reports			
Report	**Description**	**Distribution**	**Delivery Date**
ACT Student Report with *Using Your ACT Results* booklet.	Printed paper report containing College Reportable Scores	Mailed to the students at the address provided	3-8 weeks following receipt of answer documents/Online test
ACT High School Report (Student Level Score Data)	Printed paper report containing College Reportable Scores	One (1) paper copy mailed to the Director of Counseling in batches until reports are delivered.	3-8 weeks following receipt of answer documents/Online test
ACT Student Score Labels	Printed label for College Reportable scores	Two (2) printed labels per student sent to the Director of Counseling; used to place College Reportable test results on a student's high school transcript/permanent record.	3-8 weeks following receipt of answer documents/Online test Shipped with ACT High School Report (see above.)
ACT High School Check List Report	List of students for who paper reports and score labels are included in the shipment of College Reportable Score Reports.	One (1) paper copy mailed to Director of Counseling, reflecting the order in which a group of reports is shipped, alphabetically within grade in school. Checklists are not cumulative.	3-8 weeks following receipt of answer documents/Online test Shipped with ACT High School Report (see above.
ACT Student College Report(s)	College Reportable Scores are reported to students' selected colleges (up to 4)	Colleges determine frequency and format of receiving scores (e.g., paper, CD, internet)	3-8 weeks following receipt of answer documents
ACT Student Online Scores	Web page containing College Reportable Scores earned by a student from all types of testing (e.g., in school and national)	Student logs on to our student website to access a variety of services (like requesting additional college reports) through their ACT student web account.	Scores will be available online about one week after the student receives the printed score report in the mail.

2. Standard ACT In-School Reporting and Deliverables

Currently, ACT provides aggregate reports and individual data to the state and the entities they define within their jurisdiction. Reports are provided after records are processed and scores are validated.

State Contract: School-Level Deliverables		
Deliverable	**Description**	**Distribution**
ACT Profile Report– High School	An aggregate report that provides trends and averages of the High School based on the state tested grade 11 and 12 student population.	One (1) printed copy mailed to the High School Principal
State Allowed Letter (if state selects option)	A report providing student scores for an administration conducted under state allowable conditions. These scores are marked as non-college reportable.	(2) copies per student

State Contract: District-Level Distributors		
Deliverable	**Description**	**Distribution**
ACT Profile Report– District	An aggregate report that provides trends and averages of the District based on the state tested grade 11 and 12 student population.	One (1) printed copy mailed to the District Assessment Coordinator
ACT Student Level Data File – District	A Student Data File that includes College-Reportable Scores for students for whom ACT processed an answer document.	One (1) copy on CD, mailed to the District Assessment Coordinator

State Contract: State-Level Deliverables		
Deliverable	**Description**	**Distribution**
ACT Student Level Data File – State	A Student Data File that includes student scores for those whom the ACT processed an answer document.	Delivered electronically to the ADE over a secure file transfer site
ACT Profile Report–State	An aggregate report that provides trends and averages of the high school based on the state-tested grade 11 and 12 student population.	Delivered electronically to the ADE over a secure file transfer site

Reports for Custom Assessments

Arizona would have the freedom to create reporting systems for each intended audience for a custom SBCA, drawing on input from key constituent groups to deliver the kinds of state-specific measures, analyses, and performance information that matter most to Arizona parents and educators. Pearson has extensive experience developing effective custom reporting solutions for states that involve key stakeholder input, including the AIMS and AZELLA.

With a custom SBCA assessment, the ADE would also have the freedom to specify that reporting systems conform to Arizona's emerging Ed-Fi and Common Education Data Standards (CEDS) architecture, providing for more accurate and efficient data flow between the ADE, its contractor(s), and Arizona districts and schools. This would also allow the ADE to better leverage its investment in the Arizona Education Data-Driven Decision System (AzED3S) dashboards for statewide reporting, comparison, and in-depth analysis.

Requirement

d. Describe the process and timelines for scoring the assessments. Include computer-based and pencil / paper processes and timelines, as applicable.

Response

ACT Aspire

ACT Aspire's delivery model will include responses from both paper and pencil and online tests. While there are different processes for input, once the records reach the scoring step, Pearson will have one scoring program used for all models of the ACT Aspire assessments. From the scoring step and beyond, the processing and quality steps are the same for all models.

Paper and Pencil Input

Working collaboratively, Pearson and ACT Aspire have established procedures for receiving, checking in, scanning, editing, and scoring paper answer documents returned from test sites. Pearson will use standard scanning and scoring protocols, data editing standards, and mature quality control measures for processing state answer documents. Every step of the process is designed both to protect the confidentiality of the data and to verify its accuracy. The answer documents proceed through the steps outlined below.

Answer Document Receipt and Preparation

Following testing, answer documents are returned to Pearson following standard procedures. Pearson will track the return of answer documents from schools. Upon receipt of the materials, Pearson will conduct a standard compliance review, evaluate any irregularities, and perform quality assurance steps—all of which will be initiated before the answer documents are submitted for scanning.

Acclimatization of Answer Documents

Answer documents are stored in a humidity- and temperature-controlled area for up to eight hours prior to scanning. This step provides that the documents become consistent with the testing environment. Documents are not released from the acclimation area until they meet the quality standard.

Scan and Edit

After any multi-page answer documents are slit, a variety of procedures are performed to verify the data capture accuracy of the scanners used to process students' answer documents, including the following:

- Routine adjustments to the light source and other scanner settings by Pearson scanner engineers maintain scanner calibration
- We perform batch-level checking for scan track errors, a document leading edge registration check, a motion check, and a row count check.
- Diagnostic sheets, a quality control sheet, and a multiple-sheet detector are placed at the beginning of every stack within a batch.

- Answer documents include three anchor points used to detect document growth.
- Post scan, any data edit suspect errors are generated based on defined edit criteria and records that have a suspect error are flagged for review by an editor. The editor resolves flagged suspect errors based upon defined edit criteria.
- We perform additional quality control, such as random sample editor validation of scanned responses and validation of double grid and omit thresholds.

Scoring Constructed Response Items

Depending on the nature of the item, scoring students' responses to constructed response items, including technology-enhanced (TE) items, may be completed through machine-scoring, automated artificial intelligence scoring, or human hand-scoring.

As the pioneer of digital scoring within the assessment industry, Pearson owns a comprehensive series of patents covering components of digital scoring for performance assessments. Our digital scoring system has a proven record of accomplishment. We began developing a digital scoring system in 1991, and implemented digital scoring of open-ended items for the National Assessment of Educational Progress (NAEP) in 1994.

Our patented digital scoring platform has evolved over 19 years, with scoring experts and engineers collaborating to integrate quality and automation tools into the platform. The system is currently used to assign approximately 80 million scores per year.

Unlike other vendors' systems, Pearson's computer-based scoring system is lean by design—work is pulled by the scorer vs. being pushed into their queue, where it might sit idle during non-scoring hours. Our system has efficient response routing features that eliminate processing lag times while offering robust automation tools to support quality management.

With digital scoring, the scoring process is streamlined and rigorously controlled, as follows:

- Written responses are transferred from TestNav or scanned from original test booklets, and converted into an electronic format, then distributed to qualified and trained scorers.
- The digital scoring system automatically routes responses requiring second scores or resolution reads to qualified personnel.
- Digital scoring automatically prioritizes responses that need to be scored first, such as responses by high school seniors scheduled to graduate, or equating batches.
- Scores assigned to responses are automatically captured and available for review.
- Digital scoring integrates multiple processes (routing work, scoring responses, monitoring quality, and tracking progress and workflow) into a single, efficient, user-friendly system.

Our scoring platform supports our scoring experts. Our current scoring platform has provided a solution for scoring and monitoring scoring programs since 2000. The following figure includes important features and benefits of our digital scoring technology.

Pearson's Digital Scoring Offers Many Benefits		
Features	**Advantages**	**Benefits**
Data capture technology	Allows a full range of item types to be integrated into the scoring system	Students are assessed on an assortment of skills
A unique image identification that associates the online image with the actual student document	Allows easy retrieval of the original student document, if needed	Minimal impact to scoring schedule when data on an original student document needs to be clarified
Automatic routing of student responses to scorers based upon scorer qualification and project requirements	Responses are scored only by those scorers qualified for a particular item during a specific stage (first or second read)	Student data is kept secure
System capture of assigned scores, along with the scorer's identification, whether the score was a first, second, or resolution score	Scoring process is documented	Scoring processes and quality are monitored in real-time
Online backreading, which allows content staff to conduct "read behinds" of scorers' work	Accuracy of scoring can be continually monitored on both group and individual levels	Scorer feedback can target current trends and feedback can be immediate
Transparent dissemination of validity responses	Allows an objective and systematic check of scoring accuracy	Scoring directors can monitor individual and group trends on validity to determine calibration and other training needs
Dynamic feedback and automation tools	Provides immediate feedback to scorers and takes action on scorers based on performance statistics	Timely scorer quality monitoring and intervention to adjust trends or address issues
A full complement of online reliability, frequency distribution, and production reports at the item and individual scorer level	Provides statistical reports that are part of our comprehensive system for monitoring the consistency and accuracy of scoring	Progress reports at regular intervals throughout scoring process

Digital Scoring Benefits. Arizona will receive accurate, timely results with Pearson's digital scoring system.

The Pearson digital scoring system integrates multiple processes (routing work, scoring responses, monitoring quality, and tracking progress and workflow) into a single, efficient, user-friendly system, setting the industry standard for end-to-end functionality, automation, and reporting.

Quality Monitoring

In addition, Pearson's digital scoring system can provide real-time quality monitoring by producing information about individual scorers' reliability and use of the scoring scale. It can also support automated validity checking by applying a process of interspersing "validity papers" or "marker responses" with known scores into the stream of any or all scorers scoring that type of response. Since the use of such papers is invisible to scorers, the method can provide an unobtrusive means of checking scorers' accuracy against a known standard.

The result are millions of scores assigned every year by Pearson scorers, using our robust digital scoring system, meeting or exceeding the quality standards set by each customer.

Automated Scoring

Depending on specific item types, automated scoring may be used to complement human scoring for a portion of the construction response and writing items.

Pearson's Intelligent Essay Assessor (IEA) facilitates speedy delivery of accurate, valid assessments, as documented in many peer-reviewed publications. IEA has been used in a variety of assessments including the Pearson Test of English, which measures the communication skills of international students in an academic environment. Pearson's Versant automated scoring system is also used to score the speaking component of Arizona's most recent edition of the AZELLA.

IEA combines background knowledge about English in general and the subject area of the assessment in particular with prompt-specific algorithms to learn how to match student responses to human scores. IEA is adapted especially for each prompt, based on answers that actual students write in response to that prompt and on human readers' scores for those answers.

Because of the following efficiencies, Pearson's system can eliminate one to four weeks from the traditional high stakes assessment scoring cycle, while maintaining high quality standards.

Efficient Automated Scoring by Pearson	
Training	**Scoring**
IEA can be trained with fewer responses than other automatic essay assessment applications need, so it takes less time to ready our system for scoring assessments.	IEA scores a response in seconds, requiring only a fraction of the time it takes a human scorer to evaluate a response, reducing overall time for scoring.
Delivering	**Uploading**
Student answers upload directly into our system, saving the time needed to process and distribute responses to scorers.	IEA can be combined with our data warehouse, so grades can be uploaded for streamlined data management, saving administration time.

Efficient Automated Scoring for SBCA. With Pearson's automated scoring system, the ADE can save time and maintain high quality standards.

In tests of over thousands of responses, Pearson's automated scoring technology has proven as reliable as professional human scorers and more predictive of the average of two human scorers than the inter-rater reliability. IEA goes beyond measuring the grammatical correctness of a response to evaluate its content and completeness. This automated scoring technology can be used to evaluate written responses in any subject area with reliabilities equivalent to human scorers, as shown in the following figure.

Automated Scoring Performance			
Data Sets	**Number of prompts studied**	**Machine-Human Reliability**	**Human-Human Score Reliability**
Prentice Hall LA (6-12)	81	0.89	0.86
MetaMetrics	18	0.91	0.91
Higher Education	8	0.86	0.83

Automated Scoring Performance. Automated scoring can be used in many knowledge-based tasks and yields results that equal or, in some instances, surpass human scoring performance. In addition, automated scoring is unbiased and nearly instantaneous.

ACT QualityCore

Test Scoring and Resolution

Paper-and-pencil student answer documents are sent from the customer to ACT, where the demographic and multiple-choice item responses are scanned, edited, and loaded to a master database, and constructed-response items are imaged, evaluated, and scored by human readers. “Demographic” or “Non-Cognitive” student-, classroom-, school-, district- and state-level are entered in hierarchies prior to testing, enabling the retrieval of individual and aggregate demographic and “cognitive” data at those levels on a real-time basis as scores become available through an online interface.

Paper Administration (2014–2015). Multiple-choice items and constructed-response items administered on pencil-and-paper and scored by ACT.

In mathematics, science, and social studies, constructed-response tasks have three main components: the stimulus, the scoring criteria, and the holistic rubric. The stimulus is the “problem” or the task that elicits student demonstration of knowledge. It is descriptive and detailed enough that examinees know what is being asked of them. The task is focused so that it has a limited number of valid answers.
Two evaluation tools are used to properly score each task. The first is the scoring criteria identifying the explicit information (content) an examinee needs to include for a valid, thorough response to the task. Each task has scoring criteria unique to that task. The stimulus and the scoring criteria are developed in tandem to confirm that the scoring criteria contain only information elicited by the task and that the information is correct and complete.

The second evaluation tool is the holistic rubric, which is used to assess the overall proficiency of the response at four distinct levels of thoroughness. The holistic rubric assesses “big picture” criteria that are relevant to tasks regardless of specific content. For instance, in science, student responses are assessed on their ability to address content correctly and thoroughly, explain the scientific concepts and principles associated with the stimulus, and communicate their reasoning effectively and efficiently.

ACT QualityCore constructed-response tasks go through a rigorous review by external content experts. These experts confirm that the tasks accurately assess the content standards assigned to the item, are free of factual and graphical errors, are clearly presented, and assess what high school students should know and be able to do at the end of a rigorous course of study.
The holistic rubric is also rigorously reviewed by external content experts, in each discipline when items are sent out for review. This provides a check-and-balance system, maintaining continuity between items and providing that items can be assessed using the same holistic rubric.

Before items are selected for operational use, the scoring criteria are further reviewed during the field-testing process. Student responses to the constructed-response tasks are read and scored by trained ACT scorers during rangefinding. Rangefinding involves test development specialists, other content experts, and expert scorers previewing student responses to determine whether the content-specific scoring criteria for each task accurately reflect and encompass the acceptable student responses.

During rangefinding, student responses are individually scored by multiple scorers using the scoring criteria and the appropriate rubric. Responses that do not receive the same score by all scorers are discussed by the entire group and a consensus score is reached. Additions and clarifications to the scoring criteria are made at this time, if necessary. On ACT QualityCore English constructed-response tasks, which do not employ scoring criteria, discussion centers on correctly identifying those responses that best exemplify each score point on the scoring guide.

ACT QualityCore scorers are thoroughly trained before scoring operational responses. The use of well-constructed training materials is integral to the training and operational scoring process. Operational training materials include a robust anchor set and multiple practice sets for each constructed-response task. In addition, operational training materials include at least two qualification sets that prospective ACT QualityCore raters must pass in order to score operational responses.

Each response included in the training materials is analyzed by experts. An articulation (rationale) explaining why a particular score was assigned accompanies each response. The articulation explains how the holistic rubric and scoring criteria were used to determine the score. Citations from the exemplar response are included, where appropriate, to illustrate the claims made in the articulation.

The training prospective scorers receive is intensive, and when ACT QualityCore was first implemented, training was delivered on-site by test development specialists. Considerable time was spent discussing the scoring rubric so that scoring staff understand the implications of overall holistic or individual analytic domain scoring and how the rubric criteria should be applied at each score point. Also, expectations for student performance were explained, and scoring staff were shown how student performance varied at each score point.

Course rigor was stressed, and the implications for these expectations and for scoring responses were discussed in detail. Scorers learned how the rubric interacted with task-specific scoring criteria—the specific content knowledge and understanding that students must demonstrate in each response—and how both the rubric and the scoring criteria are used to evaluate each response. Scorers also learned to evaluate performance in relation to rubric criteria—not just correctness of content.

Going forward, training is being adapted based upon lessons learned and to leverage to Pearson's online training platform to promote scoring quality and consistency over time. The online training model also supports a distributed scoring system; an important part of managing the program's planned growth over time, as well as managing varying receipt patterns.

Once operational scoring commences, scorers are continuously monitored. One of the most important ways scores can be monitored is through backreading. Once scorers assign a score to a response, scoring supervisors read through the response to determine if the scores assigned are accurate and consistent with the training materials. The scoring supervisors provide written feedback and interact with scorers one-on-one to correct any discrepancies between scorers and experts. At the beginning of scoring projects, backread rates are set high so that scorers in need of remediation are promptly identified. Backread settings can be modified by scorepoint or by scorer as need arises, allowing scoring supervisors to tailor the monitoring system to the task being scored.

Scoring supervisors create recalibration sets by backreading operational responses. Recalibration sets are constructed from student responses that are typical and/or especially challenging to score. These responses are chosen during operational scoring as the scoring supervisor becomes aware of particular scoring issues that need to be addressed. Recalibration sets are given on a regular basis to address particular issues and to enable scorers to recalibrate themselves after a long break or when they switch from one item to another. Recalibration sets are administered like qualification sets; responses are discussed among supervisors and scorers. Additionally, student responses with expert scores assigned to them, known as blind validity responses, are inserted regularly into operational scoring to monitor the accuracy rate of individual scorers.

Scoring supervisors analyze group and individual scorer statistics to confirm that the scoring accuracy remains high and that scorers are maintaining high standards of efficiency and consistency. Scorers with scoring accuracy significantly lower than the standard are closely monitored by increasing the number of their scored responses that are backread by scoring supervisors. Scoring supervisors work directly with these scorers and provide feedback to confirm that their scoring accuracy improves to the standard. If improvement does not occur, scorers become ineligible to score specific items. Scorers are encouraged to ask questions and request feedback from expert scorers, as needed. Scorers unable to meet the minimum qualification requirements are released from a scoring.

As evidence of achieving quality on the ACT QualityCore scoring program, the figure below shows that Pearson has outperformed scoring expectations on every metric.

[redacted]					
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]					

Inter-rater Reliability Statistics of the End-of-Course Assessments. As shown by the inter-rater reliability metrics on current contracts, the Arizona End-of-Course assessments will continue to receive accurate scoring services.

CBT Administration (2014–2015 Administration)

ACT currently offers CBT administration for both multiple choice and constructed response only for English courses. ACT is committed to offering computer-based testing for all courses in the winter of 2014-2015. The administration of ACT QualityCore for Arizona's end of course assessments will be available to users in Arizona on this standard platform.

Report turnaround for paper-based testing is 10 days from ACT receipt of materials, provided materials are sent back without errors and students are in rosters in the system. For multiple choice CBT, tests

results are usually available almost immediately (within the hour of submitting a student's test for scoring). Multiple-choice/constructed-response CBT results are returned within 10 days, to incorporate CR scoring.

The ACT

ACT has more than a 50-year history of careful and accurate scoring and reporting. Quality assurance of scores for The ACT will be implemented to verify that students' responses have been correctly scored, and that the item scores have been properly transformed into scale scores for reporting. Generally, reports are available 3 to 8 weeks following the receipt of answer documents. Scores will be available online about one week after the student receives the printed score report in the mail.

[redacted]

Paper and Pencil Input

Working collaboratively, Pearson and ACT, Inc., have established procedures for receiving, checking in, scanning, editing, and scoring answer documents returned from test sites. Pearson will use standard scanning and scoring protocols, data editing standards, and mature quality control measures for processing state answer documents. As this is the established process for ACT answer documents, every step of the process is designed both to protect the confidentiality of the data and to verify its accuracy. The answer documents proceed through the following steps:

Answer Document Receipt and Preparation

Following testing, the answer documents will be returned to ACT following our standard procedures for in-school testing. ACT will track the return of answer documents from schools. Upon receipt of the materials, ACT will conduct a standard compliance review, evaluate any irregularities, and perform quality assurance steps—all of which will be initiated before the answer documents are submitted for scanning.

Acclimatization of Answer Documents

Answer documents are stored in a humidity- and temperature-controlled area for up to eight hours prior to scanning. This step provides that the documents become consistent with the testing environment. Documents are not released from the acclimation area until they meet the quality standard.

Scan and Edit

After any multi-page answer documents are slit, a variety of procedures are performed to verify the data capture accuracy of the scanners used to process ACT answer documents, including the following:

- Routine adjustments to the light source and other scanner settings by Pearson scanner engineers maintain scanner calibration.
- We perform batch-level checking for scan track errors, a document leading edge registration check, a motion check, and a row count check.

- Diagnostic sheets, a quality control sheet, and a multiple-sheet detector are placed at the beginning of every stack within a batch.
- Answer documents include three anchor points used to detect document growth.
- Post scan, any data edit suspect errors are generated based on defined edit criteria and records that have a suspect error are flagged for review by an editor. The editor resolves flagged suspect errors based upon defined edit criteria.
- We perform additional quality control, such as random sample editor validation of scanned responses and validation of double grid and omit thresholds.

[redacted]

Multiple Choice Scoring

Multiple choice items will be scored zero/one, with no penalty for guessing. Scoring keys and raw-to-standard score conversion tables are used to determine scores. Physical and electronic readouts of tables pass stringent quality control and sign off prior to processing.

Based on rules established in the ACT scoring program, individual answer documents may be placed on hold for review by ACT staff when the test form gridded by the examinee is inconsistent with other factors in the constant file, such as the test date, test type, and test booklet serial numbers. Such instances are resolved through hand scoring by ACT staff.

Multiple Choice Item Analysis

For each in-school test event (initial, makeup, window), we conduct an item analysis on a large sample of scored results. The analysis is conducted to verify that items perform as expected and that the scoring key has been correctly applied. No reports will be produced until the analysis is approved by ACT psychometric staff.

Additional Quality Control Measures

Our quality control includes the following additional steps:

- Diligent checking of first production records. ACT staff review the reporting results prior to public release.
- Hand scoring of statistical sample of records for score validation which are compared to reported scores.

Writing Test Performance Scoring

The ACT Writing Test is an achievement test designed to measure the core competencies of effective student writing. In designing the Writing Test, ACT examined closely the prevalent writing instruction and assessment practices of prominent secondary and postsecondary institutions across the nation. ACT's

research also included thorough review of postsecondary admission and placement exams; state writing content standards; scholarly literature on writing pedagogy and assessment; and the results of the 2002-2003 ACT National Curriculum Survey®.

As part of this development process, ACT formed the ACT National Writing Test Advisory Panel, a group composed of experts on writing instruction, writing assessment, and ESL writing. Informed by its body of research and the counsel of the Advisory Panel, ACT documented the writing skills and understandings necessary for students to write proficiently in first-year college courses. ACT drafted a scoring rubric based upon this framework. ACT and the Advisory Panel drew upon a rich body of student evidence, compiled by way of extensive field-testing, in making final refinements to the scoring criteria, ultimately producing a six-point holistic scoring rubric that serves today as the tool of student evaluation for the ACT Writing Test.

ACT employs several methods to achieve consistent application of ACT's scoring rubric. Prior to operational scoring, reader training is supported by an extensive "rangefinding" process, in which a broad sample of student responses for each operational prompt are independently rated and ultimately discussed by the rangefinding committee—a group of expert raters and ACT Writing Test developers with extensive experience with the Writing Test. The committee attaches consensus scores to individual exemplar papers, and writes accompanying score rationales for each exemplar. Then, using these materials, the rangefinding committee assembles prompt-specific training sets that effectively demonstrate the range of writing and response types that will appear in operational scoring. These training sets are used to anchor scorers to a common understanding of the rubric and its application.

The rangefinding process cultivates consensus understanding of the rubric and its evaluative criteria between test developers and scoring leaders. This tight calibration provides consistent application of the scoring rubric, irrespective of prompt, and provides a reliable measure of student performance across test administrations and calendar years.

To score operationally, readers must undergo extensive training in a baseline operational prompt and must pass certification by correctly scoring a certain percentage of papers in each of two Qualification Sets. As new prompts are introduced, readers are given additional training specific to the new prompt content and to the types of papers that have been written in response to that prompt. These up-front training and certification processes provide that students will receive accurate scores from Day 1 of scoring, and these processes pair effectively with backreading, validity paper, and calibration processes (described below) to provide a constant and reliable level of training support for readers throughout the operational scoring window. Because of this rigorous attention to consistency in development, training, and scoring, the ACT Writing Test has been a reliable predictor of college and career success for nearly a decade.

A comprehensive performance-scoring system upholds this stable scoring practice. The system has achieved the internationally recognized ISO (International Organization for Standardization) 9002 standard for quality assurance. The expertise of the scoring staff is supported by the advanced functionality of this system, which includes the following:

- **Reader performance reports** provide scoring directors with statistics on reader agreement and score distributions that allow expert backreaders to perform targeted quality assurance checks that can focus on particular readers and areas of concern. Multiple filters can be used for an immediate search of responses scored by a particular reader, responses receiving a particular score point, or a combination of both. The system also can retrieve responses where two scorers assigned scores with perfect agreement, adjacent agreement, or non-adjacent scores. Supervisors can confirm that the correct score was assigned or override an inaccurate score. Together these backreading features allow Supervisors to continuously monitor reader performance and give immediate, constructive feedback.
- **The use of validity papers**—papers previously assigned a consensus score by expert raters and then inserted into the flow of "live" operational responses—provides an objective and systematic check of scorer accuracy. The validity responses are interspersed with live responses for each scorer at a regular interval throughout the scoring day, and cumulative statistics are maintained on each reader's agreement with the validity papers' consensus scores. Poor-performing raters are redirected to focused remediation.
- **Calibration papers** are used for continued training for readers. They proactively promote accuracy by exploring prompt-specific issues, score boundaries, or types of responses that are particularly challenging to score. Online calibration sets can be launched to scorers in the field at any time. Scorers will receive a message directing them to the calibration set, and will begin their scoring shift by scoring the calibration set. Following scoring, review pages display the set again with important feedback for the reader.

Scoring a Custom Assessment

Scoring

One of the most important new features of next generation assessment is the ability to include rich, engaging content and sophisticated new item types that require students to operate at a deeper level than typical multiple-choice or fill-in-the-blank assessments. The use of technology-enhanced items (TEI), simulations, and multi-media delivery are essential for providing an online testing experience that parallels the learning experience most students will be receiving in today's next-generation classrooms.

Although most TEIs can be designed to elicit student responses that can be evaluated and scored automatically, items that allow students greater latitude in their responses, such as chemical equations, maps or diagrams, and written paragraphs or essays, often require hand scoring. To provide for deeper, more diverse measurement of students' knowledge and skills, high quality next generation assessments are designed to include a mix of these item types.

However, the added depth and rigor of next generation assessments often comes with the tradeoff of significant added time and cost for scoring, both in terms of the demand it places on contractors' scoring systems and the added labor required to hand score more open-ended items within required reporting deadlines. And when paper-based forms requiring additional processing time are added to the mix, the challenge of balancing efficiency, cost, quality, and turnaround time is compounded even further.

As previously mentioned, a custom SBCA designed by Arizona could include varying proportions of traditional selected-response, technology-enhanced, simulation, and constructed-response items designed for either machine scoring, automated artificial intelligence scoring, or human scoring. Selecting the right mix of item types and scoring methodologies can have a significant impact on turnaround time and cost. Pearson has experience designing and scoring large-scale assessments that incorporate virtually every combination of these factors, providing us with the experience it takes to design efficient, effective, and affordable next generation assessments for Arizona's SBCA.

One common approach for balancing these competing forces is to administer paper-based assessments and items that take the longest to score in an earlier administration window, with computer-based administrations of machine-scorable or AI-scored items following at a later time. This approach is similar to the one PARCC has taken for its performance-based assessments (PBA) and end-of year (EOY) assessments. The advantage of this design is that, shortly after the second administration window has closed and scoring has completed, scoring from the first administration window can be completed as well, and students' scores from both portions of the assessment can be quickly merged for reporting.

Requirement

e. Describe how scores on the assessments will be comparable to other common college/career ready assessments.

Response

ACT Aspire

Identifying a Student's Predicted Path

Parents and educators in Arizona are interested in knowing how students are proceeding toward college and career readiness. The ACT Aspire reporting will not only report growth between summative assessments, but also identify a student's predicted path and the gaps that might exist between that path and college readiness. Reporting will focus on the progress toward college and career readiness rather than a single result, highlighting the paths to improvement.

Additionally, reporting scores and scales include the following benefits:

- English, math, reading, and science will be reported using a three-digit vertical scale for grades 3–10, which is linked to the ACT test scale through the inclusion of ACT Readiness Benchmarks at each grade level
- New readiness measures within and across subjects to include ELA Scores for those who test in English, reading, and writing; reading text complexity scores, work readiness scores for those who test in reading and math
- Additionally, a STEM score will be available for schools in Arizona that elect to test with both math and science.
- Standards-based reporting will include scores in reporting categories based on the ACT College Readiness Standards and aligned to the Arizona College and Career Ready Standards.
- Progress charts showing student longitudinal growth versus the ACT Readiness Benchmark for each subject
- National norms
- ACT Aspire composite scores and predicted ACT subject/composite scores (grades 9 and 10 only)
- Classroom and aggregate reports with aggregated presentations of same student score report data in easy-to-understand visuals

ACT QualityCore

ACT QualityCore assessments have an empirical link to college performance, and therefore the percentages of students meeting the college readiness benchmark on ACT QualityCore assessments can be compared to the percentages of students meeting college readiness on other assessments.
ACT has been engaged in the discussion with other criterion assessment providers, including the consortia, around creating linkages among assessments.

The ACT

Using Your ACT Scores

Parents and educators in Arizona are interested in knowing how students are proceeding toward college and career readiness. The ACT Student Score Report will provide valuable insights into their college and career planning. Not only does the report show how they scored on the test but also provide comparisons for National Rankings.

If the student has provided college choices, summary information about each college is listed on the score report. This can help the student compare important factors about the colleges listed.

The back side of the score report assists with planning for their educations and career. If the student provides responses to the Interest Inventory questions and Student Profile Survey valuable information will be provided. The World-of-Work-Map provides insight into matching the student's interests with possible career fields.

6. Assessment Development

Requirement

a. Describe how the development of the assessments will adhere to the principles of universal design, so that the testing interface, whether paper- or technology-based, does not impede student performance.

Response

Act Aspire

Applying the Principles of Universal Design

ACT develops test materials with the broadest range of students in mind, including students with disabilities and English language learners. ACT staff members routinely evaluate test development procedures to build upon and better incorporate the principles of Universal Design and Evidence-Centered Design—already central to ACT's development work. These two frameworks, each vital to the test development process in its own right, have proven particularly powerful when used together for thoughtful item and test development focused on the construct of interest.

By applying these principles in the earliest stages of test development and continuing throughout the development process, ACT will can reduce the need for accommodations and lay the groundwork for fair and accurate measurement for students using alternate formats. Specifically, the ACT Aspire assessments address accessibility in a variety of ways, including the following:

- Using item templates designed with student accessibility in mind
- Including educators with experience working with students with disabilities and English language learners in the development and review of test materials
- Continuing to provide professional development activities and resources for staff in principles of Universal Design and Evidence-Centered Design
- Providing a range of accessibility tools and formats within the computer-based testing platform
- Incorporating APIP into ACT's item authoring system

Arizona students with physical or learning disabilities with defined access needs may be tested under special conditions or use accommodated testing materials available from ACT. Administration of the ACT Aspire assessments with accommodations is entirely at the discretion of school personnel.

ACT QualityCore

Since its inception in 1959, ACT has designed its exams for ease of use by test takers. ACT believes that the original designers of The ACT, by establishing a good test design for users, observed in advance the seven principles of Universal Design as identified by the Center for Universal Design.

In June 2002, when the National Center on Educational Outcomes, in collaboration with the Council of Chief State School Officers and the National Association of State Directors of Special Education, published its Synthesis Report 44 on Universal Design Applied to Large Scale Assessments (Thompson, Johnstone, & Thurlow, 2002), ACT reviewed its exams' design and procedures in light of that report. We found that the original designers of the ACT had succeeded in designing for a diverse audience from the beginning, and that consequently the seven principles of Universal Design have been integral to daily editorial decisions in ACT's test development of ACT QualityCore.

That report, building on seven principles, identifies "seven elements of universally designed assessments." These, with ACT's commentary regarding ACT QualityCore assessments, are as follows:

1. Inclusive Assessment Population. ACT considers the needs of a variety of students in developing the ACT QualityCore assessments. Assessments undergo rigorous reviews both in-house and from independent consultants before being added to the formative assessment pool or placed on an operational EOC test.

2. Precisely Defined Constructs. ACT QualityCore assessments are based upon empirically-derived course syllabi and course standards, and are strictly tied to the course syllabus and course standards.

3. Accessible, Non-Biased Items. ACT works with teachers to develop the assessment items, with sensitivity to a diverse examinee population. ACT QualityCore assessment items are submitted to independent reviewers for content and for fairness. The item-response data from operational administrations are analyzed for differential item functioning (DIF) through the use of the standard difference in proportion correct (STD) and the Mantel-Haenszel common odds-ratio (MH). Fewer items have been flagged by these DIF analyses than would have been expected by chance alone. The items that are flagged are analyzed to determine if any source of actual bias went undetected by the item review process. To date, no source of actual bias has been found.

4. Amenable to Accommodations. Because the ACT QualityCore assessments have been designed from the outset according to high publishing-industry standards, in keeping with curricular materials, and in consultation with in-service educators, the tests have been amenable to accommodations across the range of testing populations, conditions, and formats. We keep tests as simple and straightforward as possible, consistent with curricular requirements—and this applies equally to vocabulary, graphics, typographic design, page layout, and the interrelationships among the elements.

5. Simple, Clear, and Intuitive Instructions and Procedures. ACT QualityCore assessments comply with the Standards for Educational and Psychological Testing (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999, p. 47) with respect to test instructions that are simple, clear, consistent, and understandable and to guidelines and procedures for test administration.

6. Maximum Readability and Comprehensibility. Rather than relying on readability formulas (whose mechanical application has been shown not always to be in students' best interest), ACT relies instead on teacher involvement, expert opinion, and consultants' commentary to gauge the grade-level and content-area appropriateness of the ACT QualityCore assessments. The 10-point "Test Item Readability checklist" given on page 15 of "Universal Design Applied to Large Scale Assessments," despite minor verbal differences, is essentially a subset of the longer list of guidelines that ACT employs in its quality assurance procedures for the development and review of test materials:

"1. Students would likely have the experiences and prior knowledge necessary to understand what the question calls for."

The research-based course syllabi and course standards of the ACT QualityCore program and ACT's test development procedures provide that the ACT QualityCore assessments are linked to school curricula and reflect the diversity of students' experiences.

"2. The vocabulary is appropriate for the intended grade level."

Engaging grade-level-appropriate teachers in the test development process, together with ACT test development guidelines and extensive consultant reviews, provides for appropriate vocabulary.

"3. Sentence complexity is appropriate for the intended grade level."

ACT's selection of item writers, test guidelines, and detailed reviewing steps provide for appropriate sentence complexity.

"4. Definitions and examples are clear and understandable."

ACT so specifies in its Item Writer Guides and internal editorial standards, and internal and external test reviewers monitor this aspect of our tests.

"5. The required reasoning skills are appropriate for the students' cognitive level."

The research-based course syllabi and course standards of the ACT QualityCore program, together with ACT's test development procedures, provides that the ACT QualityCore assessments require reasoning skills appropriate for the students' cognitive level.

"6. Relationships are made clear through precise, logical connectives."

ACT editorial standards, together with the expertise of the teachers who write the tests and the ACT staff who edit them, provide that the test text is logical and that relationships are made clear. Close reviews by content experts provide that ACT meets high standards of composition and that test text is in line with current standards and usage in each area tested.

"7. Content within items is clearly organized."

ACT editorial standards, together with the expertise of the teachers who write the tests and the ACT staff who edit them, provide that that the content within items is clearly organized. Close reviews by content experts provide that ACT test items meet high standards for clarity and logic and that the items conform to high professional standards for clarity and usage in each area tested.

"8. Graphs, illustrations, and other graphic aids facilitate comprehension."

ACT provides graphic aids for the ACT QualityCore assessment stimuli and items whenever such aids are appropriate. Some of these graphic elements are adapted from published sources; others are developed especially for the tests. Graphics are tailored for the stimulus or item they illustrate and for the grade level of the examinees, and are subject to the same close review by content experts as is the text.

"9. The questions are clearly framed."

The expertise of ACT's item writers and staff and ACT's experience over decades of test development combine to frame the ACT QualityCore assessment questions clearly. ACT's staff is assisted by in-house guides to item development and by the comments and suggestions of the many experts who participate in ACT's content and fairness reviews.

"10. The content of items is of interest to the intended audience."

ACT is sensitive to the requirement that test content be interesting to students and takes care to select materials that are interesting to examinees. ACT emphasizes to item writers that test content be interesting; periodically surveys examinees to find out what test materials were the most and least successful in gaining their interest; and attends to the advice of teachers, fairness consultants, and content experts as to what topics and approaches are, and are not, appropriate for each test at each grade level.

In sum, ACT makes every effort to maintain "maximum readability and comprehensibility" in our ACT QualityCore assessments. Throughout test development, we aim to use plain language to focus every test item on the knowledge and skill being tested—that is, to avoid any material extraneous to what is being tested—in the interest of avoiding what has been termed "construct-irrelevant variance" in measurement.

7. Maximum Legibility. Universal Design Applied to Large Scale Assessments conveniently treats legibility factors under three major headings: Legible Text; Legible Graphs, Tables, and Illustrations; and Legible Response Formats. ACT has carefully weighed the many requirements of its test formats and has achieved what we believe to be an optimal balance of features. While not every feature may be ideal, the design as a whole has the optimal features for the majority of test takers. ACT provides a comprehensive manual that outlines and describes how to administer the following four available accommodations for ACT QualityCore:

- Braille
- Audio CD

- Large Print
- Reader Script

Extended-time is also available for ACT QualityCore examinees, whether testing by paper-and-pencil or the computer-based platform.

ACT QualityCore provides materials for standard and special administrations. Students with physical or learning disabilities who cannot complete the ACT QualityCore tests in the standard time limits using the standard materials may be tested, at the school's discretion, under special conditions and/or using special testing materials available from ACT. Special test materials include braille and large print (18-point) test books, large print answer sheets, audio CDs, and reader's scripts.

ACT QualityCore's educator resources work well for a wide range of cognitive abilities. Specifically, the Formative Item Pool is an assessment tool that can be used in conjunction with other ACT QualityCore resources, district curriculum tools, and interventions already in place within each district. The Formative Item Pool gives teachers the ability to assess students with varying abilities by drilling down to specific skill levels, types of questions, and course standards. Teachers can differentiate instruction by placing the cognitive level 1 questions at the front of a quiz and building the difficulty as the quiz progresses. In this way, students in need of accommodations can take the same test as everyone else in the classroom. Students under an IEP or 504, for example, would answer questions 1 to 10 only, while other students would complete questions 1 to 25. This takes the stigma out of giving students different assessments at the same time while providing a strong resource to create formative exams to assess how students are doing on specific 504 and IEP goals or to assess if an intervention is really working.

The ACT

Universal Design Accessibility

Since its inception in 1959, ACT has designed its exams for ease of use by test takers. ACT believes that the original designers of The ACT, by establishing a good test design for users, observed in advance the seven principles of Universal Design as identified by the Center for Universal Design.

In June 2002, when the National Center on Educational Outcomes, in collaboration with the Council of Chief State School Officers and the National Association of State Directors of Special Education, published its Synthesis Report 44 on Universal Design Applied to Large Scale Assessments (Thompson, Johnstone, & Thurlow, 2002), ACT reviewed its exams' design and procedures in light of that report. We found that the original designers of The ACT had succeeded in designing for a diverse audience from the beginning, and that consequently the seven principles of Universal Design have been integral to daily editorial decisions in ACT's test development.

Accessibility at ACT

ACT is committed to delivering content on its website and in its testing programs that is accessible. ACT adheres to specific guidelines and legislation, as described below, with usability as the primary standard by which it judges the accessibility of its web content and test materials.

US Rehabilitation Action Section 508

All online system content will meet Section 508 accessibility guidelines. ACT expects to be compliant with the expected revisions to Section 508 Standards within one year of their release.

Web Content Accessibility Guidelines 2.0

ACT has adopted a policy that all ACT websites and online content should achieve, at a minimum, Level A compliance with the World Wide Web Consortium's Web Content Accessibility Guidelines (WCAG), Version 2.0. This policy applies to the following:

- All public-facing informational websites
- All online applications
- Any and all corporate intranet sites

ACT is committed to achieving Level AA compliance for ACT's assessment delivery systems when possible.

Accessible Portable Item Protocol (APIP) and Personal Needs Profile (PNP)

ACT is a contributing member of IMS Global Learning Consortium and an active member of the APIP Working Group. ACT involvement in IMS work and in the APIP standards continues to increase, with ACT staff representation in the APIP Working Group and the APIP End Users group, among others. ACT is working on incorporating APIP and PNP standards into ACT's test authoring system and practice in conjunction with APIP support in ACT's delivery platforms over the next year.

Requirement

b. Describe any comparability studies between the paper/pencil and computer-based assessments.

Response

ACT Aspire

Offering Comparability Between Test Modes

ACT Aspire will launch in April 2014, providing the first longitudinal assessment to measure student progress toward college and careers. In preparation for this launch, Pearson and ACT conducted three studies in spring 2013: equating, mode comparability, and scaling. Approximately 122,000 students from grades 3 through 10 in more than 21 states participated, and 500,647 tests were administered from April to June.

In the Mode Comparability Study, grade-level tests were administered in both computer and paper versions (where most of the items were identical across the modes) to randomly equivalent groups of examinees. This allows for the examination of both item and test level information (responses to individual items, raw score distributions, option analyses, etc.) across modes. These comparisons are currently underway. An additional study is planned to confirm any adjustments needed across modes to

verify that reported scores are comparable: that is, if two students earn the same reported score, one on paper and one on computer, they are demonstrating the same level of achievement.

ACT QualityCore

ACT has conducted a comparability study which investigated the comparability of paper-and-pencil and computer-based (CBT) modes of test administration. The purpose of the study was to compare the paper-and-pencil and CBT testing modes to determine if there was a mode effect, and if so, whether the effect was significant enough to require separate conversions/adjustments to achieve comparable scores and, if necessary, to make adjustments for the first set of forms. The results of the study concluded that there was insufficient evidence to support the use of different conversions for paper-and-pencil and CBT modes of the ACT QualityCore tests.

[redacted]

Custom Assessments

Any time a test is administered in multiple modes, professional testing standards indicate the need to document the comparability of results across the modes. Pearson has conducted comparability studies for numerous custom paper and online assessments, including the AIMS and AZELLA. In addition, we conduct advanced research on the effects that different testing modes, devices, interfaces, and tools can have on student performance. Findings from these studies inform not only our own item, test, and software development activities, but we share those findings readily with our customers and colleagues through professional journals, white papers, and conferences to help advance the state of the art for the entire assessment industry. More information about our comparability and usability research can be found online at http://www.pearsonassessments.com/research.html.

If Arizona should choose to implement a custom SBCA, Pearson would consult with the ADE and its TAC to design comparability studies between paper and online-administered forms and, if desired or applicable based on the exact design chosen, between human and automated scoring.

Requirement

c. Describe the processes for item development. The description should include:

   i. How the reading and writing items will require students to demonstrate a range of higher-order, analytical thinking and performance skills in reading, writing and research based on the depth and complexity of the standards, allowing robust information to be gathered for students with varied levels of achievement; and

   ii. How the mathematics items will require students to demonstrate a range of performance based on the depth and complexity of the standards, allowing robust information to be gathered for students with varied levels of achievement.

Response

ACT Aspire

Because the Common Core State Standards (CCSS), and, by extension, Arizona's College and Career Ready Standards came largely from ACT's assessment research and data, ACT test developers have a unique and rigorous understanding of the CCSS' breadth and content. Having developed empirically-validated college and career readiness assessments for grade 8 and up over the past decades, ACT can point to student test data to confirm our domain-sampled assessments adequately cover the skills and knowledge required for college and career readiness.
(See http://www.act.org/research/researchers/reports/pdf/ACT_RR2012-2.pdf
http://media.act.org/documents/ACT_RR2012-5.pdf)

These same empirically-supported test development processes and designs are in place for ACT Aspire and ACT test development. ACT Aspire was built with a deep understanding of how students develop and demonstrate mastery as well as how they progress in achievement during their school years. ACT Aspire includes a vertically-scaled battery of achievement tests capable of measuring growth in a consistent and comparable fashion. In order to achieve this coherent, longitudinal system of measurement and provide actionable information for all students, the ACT Aspire design includes items with a range of difficulties, requiring a range of knowledge depths, and covering a progression of learning from foundational concepts to sophisticated applications.

Each test for ACT Aspire has test specifications that test developers must strictly adhere to. The test specifications describe the exact number of items covering each content area at each Depth of Knowledge (DOK) level and item characteristic to achieve sufficient depth and coverage. ACT's engagement in research serves as a first step of our test development process, functioning as part of the domain analysis step of identifying research-based knowledge and skills that students must attain to be successful in college and career. (See ACT's National Curriculum Survey describing up-to-date data about what entering college and workforce training students need to know and be able to do in college and career coursework in English language arts, mathematics, and science http://www.act.org/research-policy/national-curriculum-survey/.) Assessments are then structured to capture sufficient evidence around these claims. This process is in place for all ACT Aspire summative, interim, and formative assessments.

The ACT Aspire writing tests provide coverage of the CCSS by eliciting direct writing performance in the three text types found in the CCSS: narrative, argumentative, and informative/explanatory. The ACT Aspire rubrics developed for each grade-level assessment confirm that the breadth of the writing standards are assessed in each text type.

The treatment of text types in the suite of ACT Aspire writing assessments provides opportunities for demonstrating advanced skills and thinking. The CCSS "narrative" text type becomes Reflective Narrative in ACT Aspire, signaling that an ability to think critically about the meaning of a recounted event is an essential dimension of narrative writing competence. Similarly, ACT Aspire's Analytical Expository mode reflects the expectation that competent explanation entails analysis in the service of depth of understanding and insight. The Persuasive/Argumentative mode recognizes that good rhetorical skills include not just logos but also pathos and ethos.

The ACT Aspire rubrics (which can be reviewed in our writing exemplar materials at http://www.discoveractaspire.org/pages/test-features) are the means by which we verify that student performance is evaluated on the basis of the standards that guided task design. Each grade level has its own rubric, but there is a great deal of continuity as one moves across grades and text types. Each rubric contains four domains. The first, whether labeled "reflective narrative," "analysis," or "argument," corresponds to the Generate Ideas competency found in the Writing Competencies Model. In this domain we evaluate how well the student's ideas address the rhetorical situation, and how capable they are of meaningful reflection, explanatory analysis, or persuasive force. The "Development," "Organization," and "Language" domains in the analytic rubrics correspond to the similarly named competencies in the Writing Competencies Model.

At each grade level, and in each domain, a score of 4 is associated with "adequacy," indicating that a student who achieves this score is on track for succeeding as they enter the next grade level. At grades 6 and above, the rubrics differentiate among six performance levels; this allows for two degrees of differentiation above "adequate." A score of 5 at these grades indicates an advancing level of skill in identifying and addressing the complexities of the topic; in exploring ideas and using detailed discussions to draw out and support larger observations; in organizing with intention, aware of the effects of the sequencing of ideas; and in using strong, vivid language and effective style. A score of 6 indicates a more advanced ability in each of these areas.

The 5-point rubrics for grades 3 through 5 allow for only one degree of performance above adequate. Finer distinctions above adequate in the lower grades are not evident in the students' responses. Nevertheless, even third-graders can demonstrate higher order thinking skills by reflecting in a meaningful way on their own experiences, or those of others. The 5-point ACT Aspire rubric provides a means for recognizing that ability.

Through its task designs and rubrics, the ACT Aspire Writing test covers the breadth of the CCSS, emphasizing the place of critical thinking and the generation of ideas in proficient writing.

ACT QualityCore

Item Development and Assessment Construction Processes

Items developed for the ACT QualityCore end-of-course examinations assess content as determined by the research-based ACT QualityCore course syllabi and are aligned to the Arizona College and Career Readiness Standards and the research-based ACT Course Standards, which detail the material students should master by the end of each course. ACT QualityCore items are contextualized with scenarios that relate to students' everyday activities. The items present problem-based tasks embedded in rich, authentic contexts.

The constructed-response items require students to explain, justify, critique, create, propose, produce, design, synthesize, or otherwise demonstrate knowledge, critical thinking, and understanding in ways that cannot be assessed by multiple-choice items. These are demanding, rigorous tasks that require students to write thoughtful and detailed responses while making realistic and meaningful connections to real world

situations. When possible, context is provided to reinforce students' practical applications of concepts, theories, principles, and processes. These items will meet the ADE's item type requirements, including technology-enhanced items.

Key to the success of ACT's QualityCore research-based system is the quality of the content itself and the measurement expertise that ACT brings to both item and form development. The development of high school assessment systems has been one of ACT's core competencies throughout its 53 years of experience, and ACT has a reputation for excellence in both of these essential facets of assessment—content and measurement.

ACT's development process is designed to create a generous supply of high-quality items and valid, reliable examinations. ACT's test development procedures include:

- Periodic review and revision of exam specifications and their curriculum-based foundation
- Development, review, and evaluation of items
- Field testing of items
- Operational exam assembly

This time-tested process leverages the expertise and experience of item writers, content specialists, measurement experts, and educators to produce items for Arizona curriculum-based examinations that are valid and reliable.

Selecting and Training Item Writers

ACT's test development process is built upon decades of experience and brings together highly qualified, deeply experienced, and talented people, including item writers and content specialists.

Item writers are chosen from a pool of well qualified high school teachers who have extensive content knowledge, and who represent both genders as well as diverse ethnic backgrounds and geographic locations. These writers include active high school teachers from a variety of schools, small to large—public and private. During initial development, content specialists at ACT work closely with the writers, revising and refining content to produce high-quality items designed to meet exam specifications, reflect diversity, and meet fairness standards.

ACT provides selected-response item writers with detailed guidelines that specify the content, cognitive skill level, and format for each item; these guidelines include specifications for nondiscriminatory subject matter and language. Constructed-response item writers are given a comprehensive course specific guide for writing constructed-response items to the specifications of the ACT QualityCore program. Item writers must submit samples before they are contracted to write items that may ultimately be selected for use in field tests or on operational forms.

Field Test Review and Operational Exams Construction

When assembling operational test forms, ACT follows a systematic set of procedures designed to create examinations that fully match content, cognitive, technical, and format specifications, and that are aligned with the ACT course standards and course syllabus.

Once the EOC field test items are scored, they undergo statistical analysis focusing on difficulty, validity, and accessibility to determine whether they are suitable for operational use. Typically, items with p-values (difficulty levels) higher than .9 or lower than .2 are not assigned to an operational form and are reviewed to determine possible eligibility for future use.

The item-test score correlation (biserial coefficient) is also used to evaluate the statistical quality of items. Typically, items with biserial coefficients lower than .2 are not assigned to an operational form and are reviewed to determine possible eligibility for future use. In addition, option-response analyses are conducted. These include computing the percentage of responses for each alternative and determining the biserial coefficient of each alternative. In doing so, the statistical quality of each item can be evaluated by comparing the percentage of responses for the key and non-key options and comparing the biserial coefficients for the key and non-key options. If one of the non-key options has a higher response percentage than the key, the item will need to be reviewed carefully. Similarly, if a non-key option has a higher biserial coefficient than the key, the item will need to be reviewed carefully. Finally, differential item functioning analyses are conducted using the Mantel-Haenszel method.

Items are then chosen to fulfill the test blueprint. Operational selected-response and constructed-response exams are constructed by selecting eligible items from the item pool that align with content (blueprints) and statistical specifications. Also, a good-faith effort is made to present diversity on every exam.

The resulting operational forms are then subjected to three reviews. The first review, performed by ACT's content experts, editorial staff, and measurement staff, is intended to confirm content accuracy and match to curriculum, conformity to statistical and content test specifications, fairness, and adherence to best testing practices. The forms are then reviewed by two sets of external consultants who did not review the items prior to field-testing: content reviewers (including classroom teachers, college faculty, and curriculum specialists) and fairness reviewers. Items that do not pass a review are removed from the form; depending on the nature of the issue, such items are either discarded from the pool or are edited and placed into the development cycle anew.

The resulting ACT QualityCore end-of-course examinations are the product of evidence-centered design theory and contain a wide array of analytical, problem-solving, curriculum-based items, including in-depth constructed-response items. These assessments measure students' ability to apply the content knowledge and reasoning skills acquired in their course work to high-level items.

Scaling and Equating Methodologies

Scaling

Scaling is a process of setting up a rule of correspondence between the observed raw scores and the scale score values assigned to them for reporting purposes. The usefulness of a score scale depends on whether it can facilitate meaningful interpretations and minimize misinterpretation and unwarranted inferences (Petersen, Kolen & Hoover, 1989). The specific scaling process is described in the Technical Manual for all ACT QualityCore tests (available upon request).

The score scales for the ACT QualityCore tests were developed using the equal standard error of measurement method developed by Kolen (1988). This method was also used to establish scale scores for the ACT (Brennan, 1989), ACT Plan, and ACT Explore® tests. One of the major goals of the scaling process was to produce scale scores that would have approximately equal standard errors of measurement along most of the score scale. If this goal was met, then the same standard error of measurement could be used for all students, regardless of their scores. For most of the true scale score range, the scale score standard error of measurement was reasonably constant on all ACT QualityCore tests, at approximately 2.1.

Equating

Although each ACT QualityCore form is constructed to adhere to the same content and statistical specifications for each subject, the form difficulty may vary slightly, due to different collections of items appearing on each form. To control for this slight variation in difficulty, scores on forms within a subject are equated to the same scale so that reported scale scores have the same meaning (i.e., represent the same level of examinee achievement) regardless of the form administered. Therefore, scale scores across years for the same exam are essentially interchangeable, allowing for the monitoring of achievement over time.

So that the reported scores from the two forms are comparable, ACT QualityCore forms are equated using the equipercentile equating method (Kolen & Brannan, 2004) with postsmoothing as the primary equating method for ACT QualityCore. For the data collection, a common item nonequivalent groups design, where a set of linking items included in both the new and old forms serve, even though the forms are administered to groups of students in different years (the specific description of the equating study can be found in the Technical Manual) has been used. Recently, depending on the situation, a random equivalent groups design, where new and old forms are administered through a spiral and a single group equivalent group design, where the same examinees take all components of the ACT QualityCore (for example, selected response 1, selected response 2, and constructed response) have been used as alternative equating designs. Description of these alternative equating designs will be included in the report.

ACT has extensive experience in creating and monitoring comparable scores over time. The methodology ACT uses for equating assessments is well documented in the professional literature. In addition, stability checks on equating are routinely conducted using different methodologies, data collecting designs, and groups of examinees so that the reported scale scores are comparable across test forms and administrations. Since the SBCA is a multi-year project, it is important to compare the descriptive statistics of scale scores across years. Thus, Arizona will also receive the historical summary statistics.

Content and Fairness Review of Items

After ACT QualityCore items are developed and refined by ACT test development specialists, the items go through a rigorous review by external content experts. These experts verify that the selected-response and constructed-response items accurately assess the content standard assigned to the item, are free of factual and graphical errors, are clearly presented, and assess what students should know and be able to do at the end of a rigorous course of study. Additionally, the constructed-response scoring criteria are reviewed to confirm they are correct and complete and contain only information relevant to the item.

The materials that ACT starts with in developing ACT QualityCore assessment items reflect our nation's cultural diversity and are acquired from item writers who represent a wide variety of racial, geographic, and ethnic backgrounds. They work from detailed guidelines ACT provides that specify the test content, cognitive skill level, and item format. The guidelines also include specifications for nondiscriminatory subject matter and language. Writers are encouraged to produce materials representing diverse points of view.

Item writers whose backgrounds and experience qualify them to work on particular ACT QualityCore courses are identified. These item writers are actively engaged in teaching at the high school level and represent a variety of schools, from small private institutions to large public ones. Broad representation among item writers helps include varying perspectives and allows members of both genders and various ethnicities to participate in the development of ACT QualityCore.

Content specialists alert item writers to issues regarding the way in which various groups of the population are portrayed. It is generally recognized that unless dictated by test specifications, material likely to be less familiar to one group than to another should be excluded. Other criteria include the degree to which representatives of various groups are depicted in active versus passive circumstances, as exhibiting stereotypic mental or physical characteristics or tendencies, or as engaged in particular occupations or roles. The construction of fair assessments requires sensitivity to the changing circumstances of our society; increased variation in family structures; the multiethnic composition of the population; and a wide range of socioeconomic and urban, suburban, and rural lifestyles.

Once developed, items are reviewed by panels of external fairness experts who carefully consider each item to so that neither the language nor the content of the item will be offensive to a student, and that no item, through context or language, will disadvantage any student from any geographic, socioeconomic, or cultural background.

Item Statistics

Since the ACT QualityCore tests use classical test theory, item and test level classical statistics such as p value and point biserials will be provided. To support the bias review of items, a classical DIF analysis with the Mantel-Haenszel statistics can be computed. The yearly state reports will include Arizona-specific raw/scale score distribution summaries as well as Arizona-specific item/test statistics.

Reliability and Validity Measures

Psychometricians compute Cronbach's alpha to estimate reliability for the raw scores for the ACT QualityCore assessments. The typical range is from 0.76 to 0.94. ACT has also computed reliability estimates via the compound multinomial error model (Lee, 2007) for the scale scores. The typical range is from 0.73 to 0.93.

Content validity has been established through survey of high school teachers about the importance of the ACT QualityCore Course Standards, as well as expert review of the high school survey and course standards. Through this process, the course standards were validated or revised, as necessary. The validity argument has been enhanced through research showing that ACT QualityCore scores have a

strong correlation with concurrent course grades, are also strongly correlated with success in the next high school course (e.g., English 10 to English 11). ACT QualityCore scores are also strongly predictive of meeting the ACT College Readiness Benchmark in the same subject area.

The predictive validity evidence is summarized in *Relating End-of-Course Assessment Scores to Course Grades and College Readiness* (http://www.act.org/qualitycore/pdf/RelatingScores.pdf). ACT will continue to conduct validity studies to support different uses of ACT QualityCore scores.

Rubrics and Scoring Criteria

In mathematics, science, and social studies, constructed-response and performance tasks have three main components: the stimulus, the scoring criteria, and the holistic rubric. The stimulus is the "problem" or the task that elicits student demonstration of knowledge. It is descriptive and detailed enough that examinees know what is being asked of them. The task is focused so that it has a limited number of valid answers.

Two evaluation tools are used to properly score each task. The first is the scoring criteria identifying the explicit information (content) an examinee needs to include for a valid, thorough response to the task. Each task has scoring criteria unique to that task. The stimulus and the scoring criteria are developed in tandem so that the scoring criteria contain only information elicited by the task and that the information is correct and complete.

The second evaluation tool is the holistic rubric, which is used to assess the overall proficiency of the response at four distinct levels of thoroughness. The holistic rubric assesses "big picture" criteria that are relevant to tasks regardless of specific content. For instance, in science, student responses are assessed on their ability to address content correctly and thoroughly, explain the scientific concepts and principles associated with the stimulus, and communicate their reasoning effectively and efficiently.

ACT QualityCore EOC constructed-response tasks go through a rigorous review by external content experts. These experts verify that the tasks accurately assess the content standards assigned to the item, are free of factual and graphical errors, are clearly presented, and assess what high school students should know and be able to do at the end of a rigorous course of study.

The holistic rubric is also rigorously reviewed by external content experts, in each discipline, each time items are sent out for review. This provides a check-and-balance system, maintaining continuity between items and verifying that items in a given course can be assessed using the same holistic rubric. Before items are selected for operational use, the scoring criteria are further reviewed during the field-testing process. Student responses to the constructed-response tasks are read and scored by trained scorers during "range-finding." Range-finding involves test development specialists, other content experts, and expert scorers previewing student responses to determine whether the content-specific scoring criteria for each task accurately reflect and encompass all of the acceptable student responses.

During rangefinding, student responses are individually scored by multiple scorers using the scoring criteria and the appropriate rubric. Responses that do not receive the same score by all scorers are discussed by the entire group and a consensus score is reached. Additions and clarifications to the scoring criteria are made at this time, if necessary. On ACT QualityCore English constructed-response

tasks, which do not employ scoring criteria, discussion centers on correctly identifying those responses that best exemplify each score point on the scoring guide.

Reading Passages, Reading Indices, and Text Complexity in ACT QualityCore

As a partner in the CCSS Initiative, ACT was deeply involved in developing the standards material related to text complexity. ACT's findings about the importance of text complexity to college and career readiness, expressed most fully in the report *Reading Between the Lines* (2006), helped shape the content of the CCSS for English language arts and literacy, particularly Reading Standard 10, and the material on text complexity in CCSS Appendix A.

One goal of ACT QualityCore's rigorous program of high school instruction is that students are college- and career-ready readers by no later than the end of high school. To further that goal, ACT's QualityCore test materials are evaluated according to a pair of three-point qualitative text complexity rubrics associated with ACT's College Readiness Standards[1]. Materials meeting the criteria for *complex* texts on these scales align most closely with college and career readiness expectations as defined through examinations of common postsecondary reading materials, feedback from instructors of typical introductory postsecondary courses, curriculum surveys, and the like. These texts are meant to be accessible to college- and career-ready high school students while also reflecting the traits and demands of reading in typical first-year, credit-bearing college courses and in workforce training programs. Materials meeting the criteria for *uncomplicated* and *more challenging* texts are appropriately difficult for students making progress toward college and career readiness.

ACT's content area experts use these rubrics, along with input from experienced secondary teachers and postsecondary instructors, to help verify that reading materials in the ACT QualityCore courses are appropriately challenging for a given grade and reflect steadily increasing text complexity across grades. Given the breadth of educational attainment of the diverse populations of students who take ACT QualityCore tests at each grade, test materials will generally exhibit a range of text complexity deemed appropriate for that grade by teachers and ACT content area experts.

Cut Point Validation/Achievement Level Setting:

Arizona will receive the national user norm percentile tables, which are three-year rolling norms based on nationwide administrations. The Arizona user norm percentile tables also can be provided if requested.

Because ACT QualityCore is an existing assessment with existing predictors to college and career readiness indicators, we suggest conducting a cut point validation rather than a full achievement level setting, which will lower the overall cost of the solution while still assuring predictive validity of the assessment.

[1] See http://www.act.org/standard/planact/reading/passages.html

ACT Assessment

Test Construction

Test development at ACT is viewed as a process. The activities involved in developing a test are highly interrelated and dependent, requiring that the entire test development process—from test design through test scoring, analysis, and evaluation—be seen as a unified and interconnected system.

The test development cycle required to produce ACT assessments takes over two years and involves several stages, beginning with a review of the test specifications. Two types of test specifications are used in developing the tests: content specifications and statistical specifications.

Content specifications for ACT assessments were developed through the curricular/domain analysis discussed above in ACT Test Design. While care is taken to provide that the basic structure of the tests remains the same from year to year so the scale scores remain comparable, the specific characteristics of the test items used in each specification category are reviewed regularly. Consultant panels are convened to review the new forms of the test to verify their content accuracy and the match of the content of the tests to the content specifications. At this time, the characteristics of the items that fulfill the content specifications are also reviewed. While the general content of the test remains constant, the particular kinds of items in a specification category may change slightly.

Statistical specifications for the tests indicate the level of difficulty (proportion correct) and minimum acceptable level of discrimination (biserial correlation) of the test items to be used. The distribution of item difficulties was selected so that the tests will effectively differentiate among students who vary widely in their level of achievement.

Test materials must meet ACT's specifications for content accuracy, word count, item classification, item format, and language. During the editing process, test materials are reviewed for fair portrayal and balanced representation of groups within society and for nonsexist use of language. Materials undergo rigorous pretest analysis before being eligible for inclusion on an operational assessment. In addition, they undergo statistical analysis and further content review to confirm they adhere to ACT's rigorous standards.

The preliminary versions of the test forms are subjected to several reviews to confirm that the items are accurate and that the overall test forms are fair and conform to good test construction practice. The preliminary versions of the test forms are then submitted to content and fairness experts for external review prior to the operational administration of the test forms. These experts are not the same individuals consulted for the content and fairness reviews of tryouts. The content consultants review the forms for content accuracy, educational importance, and grade-level appropriateness. The fairness consultants are diversity experts in education who represent both genders and a variety of racial and ethnic backgrounds. The fairness consultants review the forms to help provide fairness to examinees. In all, at least 16 independent reviews are made of each test item before it appears on a national form. The many reviews are performed to help verify that each item is optimally designed to collect the best evidence of a student's level of achievement and that it is accurately and fairly evaluated.

After each operational administration, item analysis results are reviewed for any abnormality such as substantial changes in item difficulty and discrimination indices between tryout and national administrations. Examinees are encouraged to challenge any items they feel are questionable in correctness. Once a challenge to an item is raised and reported, the item is reviewed by the experts in the content area the item is assessing. In the event that a problem is found with an item, necessary actions will be taken to eliminate or minimize the influence of the problem item.

Also, after each operational administration, differential item functioning (DIF) analysis is conducted on the test data. DIF can be described as a statistical difference between the probability of the specific population subgroup (the “focal” group) getting the item right and the comparison population subgroup (the “base” group) getting the item right given that both groups have the same level of expertise with respect to the content being tested. Flagged items are reviewed by the content specialists. In the event that a problem is found with an item, necessary actions will be taken to eliminate the influence of the problem item.

ACT's entire Test Development area is comprised of a strong team of content experts--including ELA and math staff who served on the Common Core State Standards (CCSS) authoring team. They share a wealth of experience across contents assessed on the ACT tests, classroom experience, and in-depth knowledge about national and international K–12 standards. The experience of test development staff also covers alignment studies, standard-setting studies, designing customized score reports, and best practices in test development.

ACT is adept at meeting and satisfying contract deadlines. Similarly, ACT recognizes the importance of collaboration on multiple levels to deliver a successful and efficient program for customers. ACT’s work in on several key state contracts demonstrates this success. In addition, ACT has not had a contract terminated for non-performance by any customer.

Requirement

d. Describe the procedures used to ensure all test items are properly aligned to applicable standards and avoid bias. Include the role of state representatives in these processes.

Response

ACT Aspire

Alignment to Arizona Standards

The CCSS and Arizona’s College and Career Ready Standards create a shared vision and language and are based in part on ACT’s framework of college and career readiness, and developed with leadership from ACT content experts.

More than ever, educators realize they cannot let students fall behind, at any age. The new ACT Aspire program, capped by The ACT, will support teachers and students with valid, actionable information to let them know where they stand. ACT envisions not only a set of high-quality, vertically scaled summative assessments, but also a strong system that includes interim and classroom formative components to

support teachers all year long. Results presented in the common language of CCSS will help keep teachers aligned with the spirit of CCSS, as well as providing state accountability.

ACT employs an extensive program of research for assessment development that resulted in an empirically based, externally validated vertical scale measuring the construct of college and career readiness in each ACT Aspire content area. ACT's empirical understanding of the constructs allows for domain sampling in very specific ways to support accurate and reliable summative inferences of college and career readiness. Domain sampling occurs up and down the construct to verify we have adequate information to predict success. We have carefully constructed the ACT Aspire assessments to sample the domains effectively within a limited testing time. This way, students do not lose valuable instructional days when they take summative assessments.

Providing Accountability. Traditionally, alignment methodologies have been based on content alignment. This can be item-to-standard alignment or standard-to-standard alignment. Current definitions of alignment have expanded to include comparing content coverage between an assessment and other curriculum documents. Alignment in the context of No Child Left Behind (NCLB) focuses on content agreement between state curriculum standards and state assessments. There are several well-documented methodologies in support of this definition, including the methodology developed by Norman Webb.

Driven by Domain Sampling. Domain sampling, as a key component underlying ACT test blueprints, will drive the outcome of this process to provide detailed information about the alignment of ACT products to the CCSS and Arizona's College and Career Readiness Standards. With this in mind, ACT has begun conversations with leading accountability experts and is currently undertaking an independent study of the alignment of CCSS to ACT Aspire, in a form that is useful for state accountability. ACT recognizes that educational institutions must meet accountability requirements. Such requirements often dictate an independent analysis as a prerequisite for commitment to a particular assessment. ACT welcomes independent evaluations of our assessments as opportunities to discuss the goals and contents of our products.

Alignment to Arizona College and Career Readiness Standards. By choosing ACT Aspire capped by The ACT, Arizona will have assessments that meet its goals for alignment to today's standards.

Bias/Sensitivity Review Committees

ACT Aspire development follows a detailed, well documented, and standardized review process. Multiple rounds of reviews occur for each ACT test form to confirm that the items are accurate and that the overall test forms are fair and conform to good test construction practice. Additional content and fairness reviews occur prior to the operational administration of the test forms. These external reviews check the forms for content accuracy, educational importance, and grade-level appropriateness.

For the ACT Aspire assessments, ACT engages diversity experts in education who represent both genders and a variety of racial and ethnic backgrounds. The fairness consultants review the forms to help provide fairness to examinees. In all, the ACT process makes at least 16 independent reviews of each test item before it appears on an operational form. The many reviews are performed to help verify that

each item is optimally designed to collect the best evidence of a student's level of achievement and that it is accurately and fairly evaluated.

This approach enables ACT Aspire to maintain rigorous and consistent assessments across administrations while saving Arizona the expense and labor investment of meetings related to custom-developed assessments. ACT welcomes Arizona's participation as members of the fairness and sensitivity review committees in an ex-officio capacity.

ACT QualityCore

ACT works with teachers to develop the assessment items, with sensitivity to a diverse examinee population. ACT QualityCore assessment items are submitted to independent reviewers for content and for fairness. The item-response data from operational administrations are analyzed for differential item functioning (DIF) through the use of the standard difference in proportion correct (STD) and the Mantel-Haenszel common odds-ratio (MH). Fewer items have been flagged by these DIF analyses than would have been expected by chance alone. The items that are flagged are analyzed to determine if any source of actual bias went undetected by the item review process. To date, no source of actual bias has been found.

ACT Assessment

Alignment

While summative assessments are critical for state accountability, no single summative test can address standards within a comprehensively developed set of state standards, such as Arizona's College and Career Ready Standards or the CCSS. Domain sampling, as a key component underlying ACT test blueprints, will drive alignment efforts to provide detailed information about ACT products and Arizona State Standards. ACT recognizes that educational institutions must meet accountability requirements. Such requirements will often dictate an independent analysis as a prerequisite for commitment to a particular assessment. ACT welcomes independent evaluations of its assessments as opportunities to discuss the goals and contents of its products.

The ACT is designed to cover the Common Core State Standards. An internal alignment of the Arizona Academic Standards with ACT Course Standards conducted in 2010 indicated strong coverage in the English, math, science, and US history domains.

Requirement

e. Describe the procedures for ongoing assessment item development, including procedures to accommodate any future modifications to the relevant academic standards.

Response

ACT Aspire, ACT QualityCore, and The ACT

One fundamental idea underlies the development and use of ACT's assessments: The best way to determine how well prepared students are for further education and for work is to measure as directly as possible the knowledge and skills that evidence supports are needed in those settings. Accordingly, curriculum study in the content areas (English, mathematics, reading, and science) is ongoing at ACT and includes reviews of tests, curriculum guides, and national standards; a one-of-a-kind nationwide survey of educational practices and expectations; and meetings with content experts. Information from these sources helps to define a scope and sequence for each of the areas measured by ACT assessments.

The ACT National Curriculum Survey®, conducted by ACT every three to four years, continues to play a primary role in guiding the development of ACT assessments. ACT surveys thousands of elementary school, middle school/junior high school, secondary, and postsecondary teachers in English/writing, reading (including English language arts and social studies teachers), mathematics, and science for the purpose of determining what skills and knowledge are currently being taught and which are considered important for success at each grade level for college and career readiness. The survey data ACT collects about what entering college students need to know and be able to do to be ready for college-level coursework helps inform our assessments and provide that they meet the needs of college and career readiness.

Requirement

f. Describe field testing procedures and timelines for the assessments. Include information regarding field testing of Arizona students, if applicable.

Response

The success of any assessment program depends in large part on the statistical data collected on its items before they are selected for operational use. ACT uses both standalone and embedded field-test designs to collect such data for ACT Aspire items. This decision was made after weighing various psychometric considerations as well as the administrative demands on teachers and students.

The embedded field-test design offers several advantages. Examinees take the tests on-grade level and their performance on test items is not mediated by motivational issues. Specifically, by masking which items are field-test items and which are not, students cannot differentiate field-test items from operational items so that no differential performance issues related to motivation occur. Additionally, the field-test items embedded within any given ACT Aspire test form will add only a fraction of administration time (just over 10 additional minutes per student).

Due to the volume of test materials needed to support ACT Aspire, the embedded design is not practical as the lone research approach. Standalone designs will be used from time to time to bolster the ACT Aspire item pools. This will help minimize the testing burden on students during operational testing. When standalone field testing occurs in the fall, ACT recommends testing with students in the next higher grade than the test material's designation to allow complete familiarity and exposure to the grade appropriate curriculum. Understanding that voluntary participation from schools can mean low student motivation, ACT expects item mean scores and item score correlations with total test scores to be somewhat attenuated. Further, post-equating analyses are used to obtain equivalent operational scores regardless of whether items were field tested during embedded or standalone events.

When field testing, ACT recommends obtaining at least 500 students for multiple-choice items and 1000 students per constructed-response items. When items are embedded on operational tests, forms are spiraled so that students responding to any given test item are representative of the population of students taking the operational test. When samples of volunteer schools are used for standalone events, ACT strives to achieve diversity in terms of ethnic, racial, geographic, economic (using free/reduced lunch percentages as a proxy), and gender of students in public schools

Arizona students can be included in field testing for ACT Aspire if ADE adopts ACT Aspire as their solution.

ACT QualityCore

After development and review phases, ACT QualityCore items are field-tested with students from rural and urban settings, small and large schools, and public and private schools. Multiple-choice items may also be embedded in operational forms. Schools are asked to administer field test items to students who have taken the relevant course in school. The field tests are administered in the spring at the end of the course.

Student responses to the field-tested constructed-response items are read and scored by trained ACT raters. Training materials for field-testing of constructed-response test items are created during "range-finding." Range-finding involves test development specialists, other content experts, and expert raters previewing student responses to determine whether the content-specific scoring criteria for each item accurately reflect and encompass the acceptable student responses.

During range-finding, student responses are individually rated by multiple raters using the scoring criteria and the appropriate rubric. Responses that do not receive the same score by all raters are discussed by the entire group and a consensus rating is reached. Additions and clarifications to the scoring criteria can be made at this time, if necessary. On constructed-response assessments that do not employ scoring criteria, such as the ACT QualityCore English exams, discussion centers on correctly identifying which responses best exemplify each score point on the scoring guide.

Field Test Review and Operational Exams Construction

When assembling operational test forms, ACT follows a systematic set of procedures designed to confirm that each examination fully matches content, cognitive, technical, and format specifications and is aligned with the ACT Course Standards and Course Syllabus.

Once the ACT QualityCore field test items are scored, they undergo statistical analysis focusing on difficulty, validity, and accessibility to determine whether they are suitable for operational use. Typically, items with p-values (difficulty levels) higher than .9 or lower than .2 are not assigned to an operational form and are reviewed to determine possible eligibility for future use. The item-test score correlation (biserial coefficient) is also used to evaluate the statistical quality of items. Typically, items with biserial coefficients lower than .2 are not assigned to an operational form and are reviewed to determine possible eligibility for future use. In addition, option-response analyses are conducted. These include computing the percentage of responses for each alternative and determining the biserial coefficient of each alternative. In doing so, the statistical quality of each item can be evaluated by comparing the percentage of responses for the key and non-key options and comparing the biserial coefficients for the key and non-key options. If one of the non-key options has a higher response percentage than the key, the item will need to be reviewed carefully. Similarly, if a non-key option has a higher biserial coefficient than the key, the item will need to be reviewed carefully.

Items are then chosen to fulfill the test blueprint, which is aligned with the ACT Course Standards. Operational multiple-choice and constructed-response exams are constructed by selecting eligible items from the item pool that align with content (blueprints) and statistical specifications. Also, a good-faith effort is made to present diversity on every exam. The resulting operational forms are then subjected to three reviews. The first review, performed by ACT's content experts, editorial staff, and measurement staff, confirms content accuracy and match to curriculum, conformity to statistical and content test specifications, fairness, and adherence to best testing practices. The forms are then reviewed by two sets of external consultants who did not review the items prior to field-testing: content reviewers (including classroom teachers, college faculty, and curriculum specialists) and fairness reviewers. Items that do not pass a review are removed from the form; depending on the nature of the issue, such items are either discarded from the pool or are edited and placed into the development cycle anew.

The resulting ACT Quality Core EOC examinations are the product of evidence-centered design theory and contain a wide array of analytical, problem-solving, curriculum-based items, including in-depth constructed-response items. These assessments measure students' ability to apply the content knowledge and reasoning skills acquired in their course work to high-level tasks.

The ACT

Item Tryouts

Items that are judged to be acceptable in the review process are assembled into tryout units for pretesting on samples from the national examinee population. These samples are carefully selected to be representative of the total examinee population. Each sample is administered a tryout unit from one of the four academic areas covered by the ACT tests. The time limits for the tryout units permit the majority of students to respond to all items.

Item Analysis of Tryout Units

Item analyses are performed on the tryout units. For a given unit the sample is divided into low-,

medium-, and high-performing groups by the individuals' scores on The ACT in the same content area (taken at the same time as the tryout unit). The cutoff scores for the three groups are the 27th and the 73rd percentile points in the distribution of those scores. These percentile points maximize the critical ratio of the difference between the mean scores of the upper and lower groups, assuming that the standard error of measurement in each group is the same and that the scores for the entire examinee population are normally distributed (Millman & Greene, 1989).

Proportions of students in each of the groups correctly answering each tryout item are tabulated, as well as the proportion in each group selecting each of the incorrect options. Biserial and point-biserial correlation coefficients between each item score (correct/ incorrect) and the total score on the corresponding test of the regular (national) test form are also computed. Item analyses serve to identify statistically effective test items. Items that are either too difficult or too easy, and items that fail to discriminate between students of high and low educational achievement as measured by their corresponding ACT test scores, are eliminated or revised for future item tryouts. The biserial and point biserial correlation coefficients, as well as the differences between proportions of students answering the item correctly in each of the three groups, are used as indices of the discriminating power of the tryout items. Each item is reviewed following the item analysis.

ACT staff members scrutinize items flagged for statistical reasons to identify possible problems. Some items are revised and placed in new tryout units following further review. The review process also provides feedback that helps decrease the incidence of poor quality items in the future.

7. Assessment Cost

On the following pages we have provided costing information for this response.

ACT Aspire State Summative and Interim Pricing

State Early Adopter Pricing • Pricing valid until December 31, 2013 for signed contracts or POs received	
Computer Based Testing	**Paper and Pencil**
$11.70 per student full battery	$18.70 per student full battery

State Summative Pricing • Post January 1, 2014	
Computer Based Testing	**Paper and Pencil**
$21.00 per student for one subject $1 additional per subject	$27.00 per student for one subject $1 additional per subject

State Formative Pricing • Effective November 1, 2013—Formative Assessment available August 2014	
Computer Based Testing	**Price per student**
25,000 to 50,000 Students	$ 9.00
50,0000 to 100,000 Students	$ 8.50
100,000 to 250,000 Students	$ 8.00
250,000 to 500,000 Students	$ 7.00
Over 500,000 Students	$ 6.00

Summary

- States are eligible for two years of early adopter pricing regardless of whether testing begins in spring or fall of 2014.
- To participate in the formative assessment for 2014–15 school year, a state must be enrolled in the summative program.

State-level Qualification Requirements for Early Adopter Pricing:

- The Early Adopter Summative Assessment Order Form must be submitted by December 31, 2013 and a P.O. received by January 31, 2014.
- States are not required to participate in the spring 2014 testing window(s) but may at their option.

ACT QualityCore Assessments

End of Course Pricing – Per Test • Multiple Choice and Constructed Response		
Subject Tested	**Paper**	**Online**
Algebra I	$28.00	$22.00
Algebra II	$28.00	$22.00
Geometry	$28.00	$22.00
English 9	$28.00	$22.00
English 11	$28.00	$22.00
English 10	$28.00	$22.00

- ACT will increase paper price $1.00 to $29.00 in 2015-16.
- ACT will increase online price $1.00 to $23.00 in 2015-16.
- ACT will offer pretests with item response online starting in spring 2015
- ACT will offer constructed response online in the next platform; however, constructed response for math courses will not be available until 2015-16.

Assumptions – deliverables

- 6 courses English 9–11, Algebra I, Algebra II, and Geometry
- Expecting 50% CBT / 50% P&P in 2014–15 with 10% movement each year toward 100% CBT

ACT QualityCore End-of-Course Assessments measure the learning outcomes all students must attain in order to succeed in college and careers. Each assessment includes problem-based questions embedded in both academic and real-world contexts that are accessible and relevant to high school students. These real-world problems require practical applications of concepts, theories, principles, and process. Flexible options include:

- Two 45-minute test sessions—multiple choice and constructed response, or multiple choice only
- Computer-based or paper-and-pencil format
- Standards-based reporting

Cost estimates presented above are based upon projected fees for assessments, training, and standard reporting. Professional development is purchased separately. Other services and requests would be negotiated between ACT, Pearson, and ADE based on the complexity of the project.

The ACT Statewide Testing Program

ACT Pricing – Per Student • Multiple Choice and Constructed Response	
Per student price for the spring 2014 Statewide ACT Test with Writing:	$48.50
Per student price for the spring 2014 Statewide ACT Test without Writing:	$33.00

ACT is devoted to helping people achieve education and workplace success. Established in 2001, ACT's Statewide Testing Program forwards this mission by offering states the opportunity to fund The ACT for their state's entire junior class. Beginning in the fall of 2014, states could also choose to fund The ACT for their state's entire senior class. The fall date will allow seniors to receive their ACT scores in time to meet normal college application deadlines.

Students participating in the ACT Statewide Testing Program:

- Are tested on what they actually learn in school
- Obtain valuable information about their academic strengths and weaknesses
- Can measure their college readiness by comparing their subject scores with the ACT College Readiness Benchmarks
- Receive college admissions scores that are accepted by all four-year colleges and universities in the United States
- Take the test during a regular school day and do not have to travel to a test site on the weekend
- Experience less stress in a familiar environment surrounded by familiar faces;
- May request scores to be sent to as many as four colleges and universities when completing the noncognitive portion of their answer document during the pretest session
- Get personalized information to make future college and career decisions
- May choose to provide information that may help identify opportunities for college admissions and scholarships

States who participate in the ACT Statewide Testing Program:

- Introduce all students to the research-based ACT College Readiness Standards and Benchmarks
- Promote a college-going culture among all students
- Encourage all students to seriously consider their postsecondary options

- Receive statewide data to assess student academic achievement and college readiness levels, inform curriculum decisions, and advise students

ACT Statewide Testing Program Pricing:

- Per student price for the spring 2014 Statewide ACT Test without Writing: $33.00
- Per student price for the spring 2014 Statewide ACT Test with Writing: $48.50

Costs for a Custom Assessment

Costs for a custom assessment would depend greatly on the scope of work included in the RFP and how certain costs were shared or handled. For example, costs for licensing items from either of the national assessment consortia could be borne by the ADE through direct negotiations with PARCC and/or SBAC, rather than being included in the vendor's contract.

One potential way to estimate costs for a custom assessment would be based on the projected amount of hand scoring for a custom blueprint versus an established baseline reference, such as PARCC.

However, reductions in test blueprints would not necessarily result in one-for-one reductions in cost. For example, a 25 percent reduction in the number of hand scored items might result in only a 5–10 percent reduction in total cost per student, depending on the nature of student responses and complexity involved in scoring the items that remain versus those removed. On the other hand, eliminating a single extended-response item, such as an essay, could reduce cost by 2–5 percent alone, depending again on the complexity of scoring. These savings would have to be balanced against the loss of content coverage and information provided about student performance.

Assumptions related to the number of paper versus online tests administered each year will also have a bearing on overall cost estimates. Based on our experience in other states, for Arizona we have conservatively estimated that 50 percent of all tests might be administered online in 2014–15, with an additional 10 percent online each year thereafter, reaching nearly 100 percent computer-based administration by 2020.

As paper volumes reach a certain minimum, other cost-saving measures such as local scoring or data entry may be feasible, similar to the approach currently taken for low-volume AZELLA placement testing. Generally speaking, depending on the number of students, the nature of the test content, and length of the test, paper-based administration will cost between $5 to $10 more per student than an equivalent online form.