Response To

Request for Information

ASTF0001

Iowa Testing Programs

November 20, 2013

THE UNIVERSITY OF IOWA

OFFICE OF THE PRESIDENT
101 Jessup Hall
Iowa City, Iowa 52242-1316
319-335-3549
Fax 319-335-0807

November 18, 2013

W. David Tilly, Ph.D.
Deputy Director and Administrator
Iowa Department of Education
Grimes State Office Building
400 East 14th Street
Des Moines, Iowa 50319

Dear Dr. Tilly:

I am pleased to be writing you in support of the efforts of Iowa Testing Programs and the UI College of Education to provide the Iowa Department of Education with up-to-date information about the assessments they have under development, the *Next Generation Iowa Assessments*. These assessments are directly responsive to House File 215 passed by the 2013 General Assembly and signed by the Governor on June 3, 2013. The University of Iowa is grateful for the opportunity to support Iowa schools and the Department with assessments that satisfy all of the provisions of HF215. The outreach and service provided by ITP in implementing these assessments will continue an 80-year legacy of collaboration between the University of Iowa and the State's K-12 education system.

These new assessments maximize alignment with the Iowa Core as amended in 2010 to incorporate the Common Core State Standards. They incorporate the latest developments in technology for large-scale assessment and in psychometric methods in tracking progress toward college or career readiness. The information they provide is bolstered by a long-standing program of research at the university on technical quality and test validation for the specific purposes of informing instruction and enhancing teacher quality and student learning. These are key aspects of reliable, valid, and fair measures as required by HF215.

The ITP response to Request for Information ASTF0001, "Educational Assessments for State Accountability," addresses the general requirements in Section 2 and provides a structure likely to generate cost savings and program efficiencies for the State.

I hope that you will give this response your full consideration.

Sincerely,

Sally Mason
President

Executive Summary

Iowa Testing Programs applauds the state's commitment to high-quality assessments for the measurement of student achievement on the Iowa Core. Since the adoption of the Common Core State Standards in 2010, Iowa Testing Programs research has been actively developing and validating the *Next Generation Iowa Assessments*. This new assessment is slated for introduction in 2015 and will offer a fully-aligned, valid, reliable and research-based assessment for consideration by the state of Iowa. This assessment introduces new item types and grade-specific content coverage to better address the rigor of the college- and career-ready standards of the Iowa Core. This assessment will also produce standards-based reporting (both electronic and paper) to help inform instruction.

As specified in Request for Information (ASTF0001) issued by the Iowa Department of Education on November 6, 2013, this document addresses the general requirements articulated in Section 2. As the Assessment Task Force considers this and other responses to the RFI, please note that the *Next Generation Iowa Assessments* offers advantages unique for the state of Iowa, including the extensive involvement of Iowa teachers, administrators and students in all steps of the development process. Highlights of this response include the following:

Aligned and Rigorous	• Assesses reading, language, writing, mathematics, science and social science standards of the Iowa Core • Requires higher-order, analytical thinking skills • Offers a range of item types to measure the content and rigor of the Iowa Core
Sustainable	• Designed and developed by ITP, a research center within the College of Education at the University of Iowa, at no cost to the state or districts in Iowa • Externally supported through ITP, not dependent upon other sources of funding for sustainability
Credible	• Written and reviewed by Iowa educators • Pilot tested on Iowa students • Validated in higher education institutions in the state of Iowa
Technically sound	• Reliability and validity evidence captured on students in the state of Iowa • Piloted online system with technical documentation and support

Iowa Testing Programs encourages questions concerning the information shared in this response and welcomes the opportunity to provide the Task Force with additional information. Iowa Testing Programs remains committed to ensuring good measurement for the set of core indicators required by Iowa Code [12.8(3)].

Table of Contents

Section 2.1 – Company Information

This response is being submitted by:

Iowa Testing Programs
College of Education
University of Iowa
334 Lindquist Center South
Iowa City IA 52242

All inquiries concerning this response should be directed to:

Stephen B. Dunbar
Director, Iowa Testing Programs
College of Education
The University of Iowa
334 Lindquist Center South
Iowa City IA 52242
steve-dunbar@uiowa.edu
Office: (319) 335-5561
Mobile: (319) 383-8125
Fax: (319) 335-6038

Section 2.2A – Information Requested

Options to Generate Cost Savings and Efficiencies

Iowa's adoption of the Common Core State Standards in July 2010 offered our educators the opportunity to transform instruction and increase student performance within the state. In support of this effort, Iowa Testing Programs (ITP) has devoted the last three years to leverage our own funding to design and develop new and innovative assessments that address these new standards.

As new assessments are designed, developed and delivered between now and 2015, ITP can provide a technically sound, aligned, valid and reliable assessment that will offer the state of Iowa savings and program efficiencies unmatched by either of the two consortia or competing test vendors. Discussed in greater detail in Section 2.2B, ITP is in the unique position of offering all design, development, psychometrics and research to the state of Iowa at no cost. This support is available due to external sources of funding that support ITP's work. No other state is in a similar position to benefit from ongoing research and development conducted by their assessment provider. The cost savings and program efficiencies present themselves through the experience and expertise of ITP personnel and outreach opportunities within the state.

Experience

ITP is prepared and qualified to help the state of Iowa respond to the ambitious call to develop assessment and reporting tools that provide information on the standards, curriculum, assessment and instruction. Beyond designing assessments, we conduct educational research and analysis to inform the utility of our assessments. As part of a university that is dedicated to advancing quality and equity in education for all learners, ITP stands ready to offer expertise, assistance and guidance to help contribute to the state's success in implementing a new assessment program as required by HF215.

Expertise of Personnel

States must be confident that their assessment systems are of sufficient technical quality, as well as support the broad range of uses of results that withstand legal challenges — particularly when used for high-stakes purposes associated with school accountability and teacher effectiveness. One significant source of cost savings and program efficiencies for the state of Iowa is generated by the dedication of university faculty, psychometricians, research scientists, assessment developers and program managers that are well-versed in the state of Iowa's needs. ITP's team of resources offers substantial research capabilities to drive innovation in assessment and advance education and equity for all students in the state of Iowa, and is in a unique position of being able to be responsive to the state's needs as those needs evolve to meet the long-term goals of education reform in Iowa.

Outreach

A second option that would generate costs savings and program efficiencies is the dedicated outreach that ITP provides as it delivers assessment results. ITP supports use of assessment information by local, regional and state audiences through workshops and professional development opportunities. This is an extremely important component of a successful assessment program is assisting in the interpretation of the results. Providing in-service activities at the district, AEA or state level helps ensure that Iowa educators see the connection between standards and assessments while using the results to inform instruction for future years.

All of ITP's in-service activities are performed as part of the public engagement and outreach mission of the College of Education Strategic Plan (2013-2018) to create professional development opportunities. That plan includes the dimensions of Student Success, Knowledge and Practice, New Frontiers, and Better Futures for Iowans as focal points. It represents a continued commitment by the university and ongoing cost savings and sustainability result from it.

Section 2.2B – Information Requested

Alignment to Iowa Core Standards

Alignment to the Iowa Core Standards has been a guiding principle of the development of the *Next Generation Iowa Assessments (NGIA).* Since the Iowa Core Standards were adopted by the state in July 2010, Iowa Testing Programs (ITP) has depended upon these standards to define and shape the development and research necessary to build an assessment aligned to the Iowa Core in English Language Arts, Mathematics, Science and Social Studies.

To produce items that are aligned, ITP follows a well-defined development process that helps to ensure the appropriate balance and representation of content. This process includes the following steps:

- Creation of test specifications that define the content areas and cognitive processes to be measured by the *NGIA*
- Development of test materials (items and scoring rubrics) by Iowa educators that measure critical aspects of the Iowa Core
- Alignment of individual items to the Iowa Core by Iowa educators during the item development process
- Verification of these alignments by focus groups of Iowa educators who are actively teaching English Language Arts, Mathematics, Science, and Social Studies at the appropriate grade levels

ITP believes that calling on the expertise of Iowa educators from the very beginning of our development process will be a defining feature of our ability to demonstrate alignment to the Iowa Core Standards.

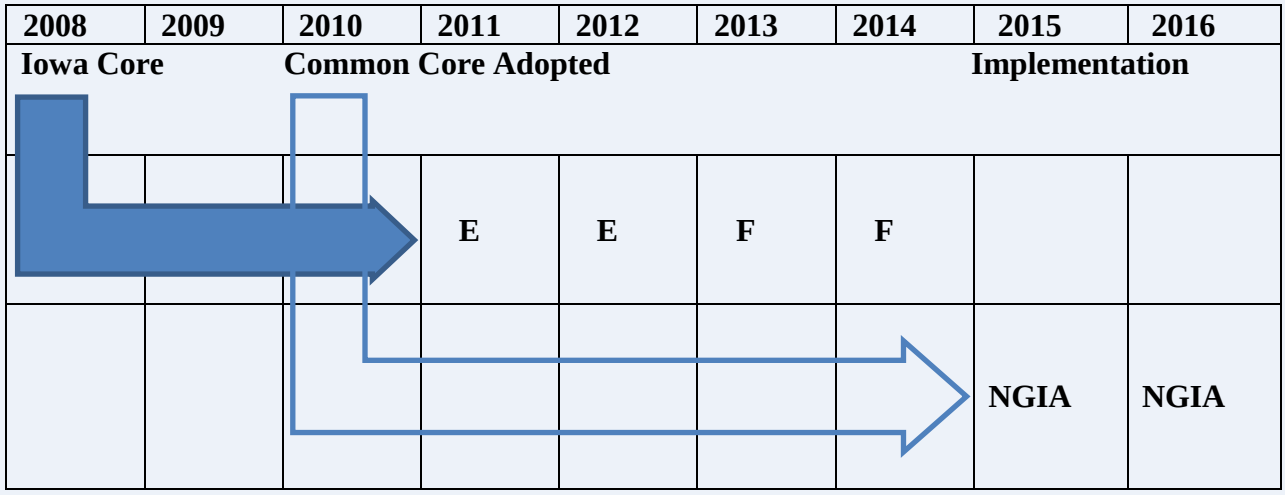
Traditionally, alignment methodologies have been based on content and cognitive definitions by examining an item-to-standard alignment. However, as with any assessment, this process cannot be completed until the item pool or forms have been completed. ITP encourages the Iowa Department of Education to commission a study to be conducted in 2015 to validate the alignment of the assessment recommended by the Assessment Task Force to the Iowa Core. New and traditional methodologies that address the complex challenge of aligning assessments to core standards will appropriately validate the coverage and rigor of the core standards and the claims and targets defined by assessments.

Clarification Concerning Alignment Study of Form E

The Iowa Department of Education recently released a report commissioned in March 2013 to evaluate the alignment of the *Iowa Assessments*, Form E, to the Iowa Core/Common Core State Standards in Reading and Mathematics. As indicated in the figure below, Form E was built to the Iowa Core as the standards were mandated in 2008, prior to the development and adoption of the Common Core State Standards (CCSS). The 2008 version of the Iowa Core was based on grade spans (primary, intermediate, middle and high school). Form E of the *Iowa Assessments* (and a parallel Form F) were built as a transition to serve as a bridge between the adoption of the Iowa Core in 2008 and the full implementation of the Iowa Core in 2015, while continuing to meet the reporting requirements of NCLB. This transition was a necessary step for the state of Iowa to measure the Iowa Core of 2008 while reporting AYP and proficiency information to the federal government. This decision was discussed explicitly with the Iowa Department of Education as plans for the transition were made.

The commissioned report explicitly states that the *Iowa Assessments*, Form E, were not built to the standards that were used to complete the study. Appendix A cites text excerpted from this report (pages 4 and 15) that refutes any findings in this report for purposes of alignment to the Common Core adopted by Iowa in 2010.

In response to the misleading claims made by this report, ITP has published a *Response to Claims Raised by DRC's Mathematics and Reading Alignment Study* at http://itp.education.uiowa.edu. We encourage Assessment Task Force members to review this document. In addition to improperly aligning Form E, apparent flaws in the methodology and implementation of the alignment study raise serious reservations about the integrity and accuracy of the results.



Test Specifications for the *Next Generation Iowa Assessments*

The test specifications provide the "blueprint" for test construction, defining the necessary steps and procedures for development. Tables 1-3 below list the domains in the Iowa Core that will be assessed by the English Language Arts and Mathematics tests of the *Next Generation Iowa Assessments* at Grades 3–8 and in high school.

Table 1. Iowa Core ELA/Literacy Domains Assessed

Iowa Core ELA/Literacy Domains	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Foundational Skills	●					
Key Ideas and Details	●	●	●	●	●	●
Craft and Structure	●	●	●	●	●	●
Integration of Knowledge and Ideas	●	●	●	●	●	●
Conventions of Standard English / Knowledge of Language	●	●	●	●	●	●
Vocabulary Acquisition and Usage	●	●	●	●	●	●
	Text includes Informational and Literature			Text includes Informational, Literature, Science and Social Studies		

Table 2. Iowa Core Math Domains Assessed

Iowa Core Math Domains	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Operations and Algebraic Thinking	●	●	●			
Number and Operations in Base Ten	●	●	●			
Number and Operations – Fractions	●	●	●			
Measurement and Data	●	●	●			
Geometry	●	●	●	●	●	●
Ratios and Proportional Relationships				●	●	
The Number System				●	●	●
Expressions and Equations				●	●	●
Statistics and Probability				●	●	●
Functions						●

Table 3. Iowa Core Domains Assessed at End-of-High School Grades

Iowa Core ELA/Literacy Domains		Iowa Core Math Domains	
Key Ideas and Details*	●	Geometry	●
Craft and Structure*	●	Statistics and Probability	●
Integration of Knowledge and Ideas*	●	Functions	●
Conventions of Standard English / Knowledge of Language	●	Algebra	●
Vocabulary Acquisition and Usage	●	Numbers and Quantity	●
*Text includes Informational, Literature, Science and Social Studies		Modeling	●

Item Types on the *Next Generation Iowa Assessments*

Measuring the depth and breadth of the current Iowa Core Standards (2010) requires a balanced and layered approach that incorporates a range of tasks and stimulus materials. Multiple-choice items, such as those used in the current *Iowa Assessments*, are excellent for evaluating student knowledge and understanding of a variety of concepts and content included within the Iowa Core. However, additional assessment formats are needed to measure those skills that are not easily assessed by these more traditional formats. The intent of increasing the types of item formats in the assessments is to expand and improve the measurement of student understanding and proficiency overall.

The *Next Generation Iowa Assessments* are designed to mirror the rigor of the current Iowa Core by employing a robust suite of traditional and nontraditional item types, including:

- **Constructed- and extended-response items:** These items challenge students to draw upon higher-order thinking and cognitive processes. For example they may require the student to solve multistep mathematics problems or craft an extended response to a writing prompt. A mix of technology-based scoring (automated scoring engines) and human scoring using scoring rubrics will be used.
- **Technology-enhanced items:** Typically administered on a computer, these items require students to make use of complex thought processes and responses. By taking advantage of the many features in today's computer-based technologies, these items can be interactive. They may include unique response interfaces such as hot spots, drag-and-drop, point-and-click, fill-in-the-blank, and graphing; or require students to provide or select multiple responses to a single question. These items are machine scored.

- **Multiple-choice:** These items are efficient to administer and offer strong technical properties. As evidenced by the *Iowa Assessments*, these items can be written to address varying levels of cognitive complexity. The multiple-choice items in the *Next Generation Iowa Assessments* measure students' skills and knowledge at three cognitive levels. This item type will serve as part of the summative assessment that is specifically designed to efficiently measure student learning in terms of the Iowa Core.

The assessments ITP proposes for the state of Iowa's use are fixed-form assessments for ELA (including Reading, Language and Writing), Mathematics, Science, and Social Studies, intended for use within the last 12 weeks of the academic year. These summative assessments will measure growth and proficiency across the Iowa Core Standards, specifically defined for each grade level. The *NGIA* will not contain any overlapping items across grade levels as has been true of previous forms of the *Iowa Assessments*.

ELA Assessments

The ELA assessments will address reading, language and writing. The following tables present a blueprint for the reading, language and writing tests within the ELA assessment. Note that the number of items listed below is approximate; the final number of each type of item will be determined as the tests are finalized and as budgetary and other considerations are evaluated by the State.

In the reading and language tests, text complexity will be addressed for each passage through quantitative and qualitative measures. Quantitative measures are aspects of text complexity that are unlikely to be evaluated by a subject matter expert reliably, and therefore computer software is used. The quantitative measure relevant to CCSS passage development is the Lexile® score. The University of Iowa follows the guidelines for Lexile ranges found on the Common Core State Standards Initiative's website. Qualitative measures are best determined by a subject matter expert who can evaluate the use, organization, language appropriateness, and the likely understanding of the target reader. The qualitative measures relevant to CCSS passage development are documented in a passage review checklist and evaluated by a minimum of two independent subject matter experts during the development process.

Table 4. ELA Blueprint—Section 1, Reading

Grade	Multiple-Choice Items	Technology-Enhanced Items	Constructed-Response Items
3	29-31	2-4	3-5
4	30-32	2-4	3-5
5	31-33	2-4	3-5
6	32-34	2-4	3-5
7	33-35	2-4	3-5
8	34-36	2-4	3-5
9	29-31	2-4	3-5
10	29-31	2-4	3-5
11	29-31	2-4	3-5

Table 5. ELA Blueprint—Section 2, Language

Grade	Multiple-Choice Items	Technology-Enhanced Items
3	28-30	5-7
4	31-33	5-7
5	33-35	5-7
6	35-37	6-8
7	36-38	7-9
8	38-40	8-10
9	44-46	8-10
10	44-46	8-10
11	44-46	8-10

The third section of the ELA assessment is the writing test, which will include a single writing prompt. The responses will be scored by two independent scorers, with the option for a third adjudication reader. The scores assigned will be summed, so the writing test can yield a maximum score of 12 points.

Table 6. ELA Blueprint—Section 3, Writing

Grade	Extended Constructed-Response Items	Scale
3	Persuasive, Narrative	12 points
4	Explanatory, Narrative	12 points
5	Persuasive, Narrative	12 points
6	Explanatory, Persuasive	12 points
7	Explanatory, Narrative	12 points
8	Explanatory, Persuasive	12 points
9	Explanatory, Persuasive	12 points
10	Explanatory, Persuasive	12 points
11	Explanatory, Persuasive	12 points

Mathematics Assessments

The *Next Generation Iowa Assessments* in mathematics will be rigorous, assessing what students can do with what they have learned. Items included in the assessment will be carefully selected from the full range of content of the Iowa Core, and will require a range of cognitive skills. Students will be required to demonstrate their understanding of concepts and procedures, to solve problems, analyze data and communicate their results.

At the high school level, the Mathematics assessments are being designed to assess the indicators as they are organized in Appendix A of the CCSS for Math, found at: http://www.corestandards.org/assets/CCSSI_Mathematics_Appendix_A.pdf. Iowa Testing Programs is developing and expanding end-of-course assessments in Algebra I, II, and Geometry, which will align with the Iowa Core Standards.

The following tables present blueprints for the Mathematics assessments. Note that the number of items listed below is approximate; the final number of each type of item will be determined as the tests are finalized and as budgetary and other considerations are evaluated by the state.

Table 7. Mathematics Blueprint: Grades 3–8

Grade/Subject Area Test	Multiple-Choice Items	Technology-Enhanced Items	Constructed-Response Items
Grade 3	37	4	3-5
Grade 4	42	4	3-5
Grade 5	47	4	3-5
Grade 6	51	5	3-5
Grade 7	56	5	3-5
Grade 8	61	5	3-5

Table 8. Mathematics Blueprint: High School End-of-Course Tests

Grade/Subject Area Test	Multiple-Choice Items	Technology-Enhanced Items	Constructed-Response Items
Algebra I	20-25	5-8	6-8
Algebra II	20-25	5-8	6-8
Geometry	20-25	5-8	6-8

Valid, Reliable and Piloted in Iowa

Iowa Testing Programs at the University of Iowa has a proud history of developing assessments that are valid, reliable and accurately measure student progress and achievement. This tradition has been and is still strongly supported by Iowa educators who serve as writers and reviewers and Iowa students who serve as pilot study participants. This process addresses the criterion of validity broadly as described in what follows.

Validity is an attribute of information from tests that, according to the *Standards for Educational and Psychological Testing*, "refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (1999, p. 9). Assessment information is not considered valid or invalid in any absolute sense. Rather, the information may be valid for a particular use or interpretation and invalid for another. The *Standards* further state that validation involves the accumulation of evidence to support the proposed score interpretations. This part of our response provides an overview of the data collected over the history of the *Iowa Assessments* that pertains to validity.

Validity must be judged in relation to purpose. Different purposes may call for tests built to different specifications. For example, a test intended to determine whether students have reached a performance standard in a local district is unlikely to have much validity for measuring differences in progress toward individually determined goals. Similarly, a testing program designed primarily to answer "accountability" questions may not be the best program for stimulating differential instruction and creative teaching.

Content Validity

The procedures used to develop and revise test materials and interpretive information lay the foundation for test validity. Inferences based on test scores, not to mention desirable consequences from those inferences, can have utility only if test development produces meaningful test materials. Content quality is thus the essence of arguments for test validity (Linn, Baker & Dunbar, 1991).

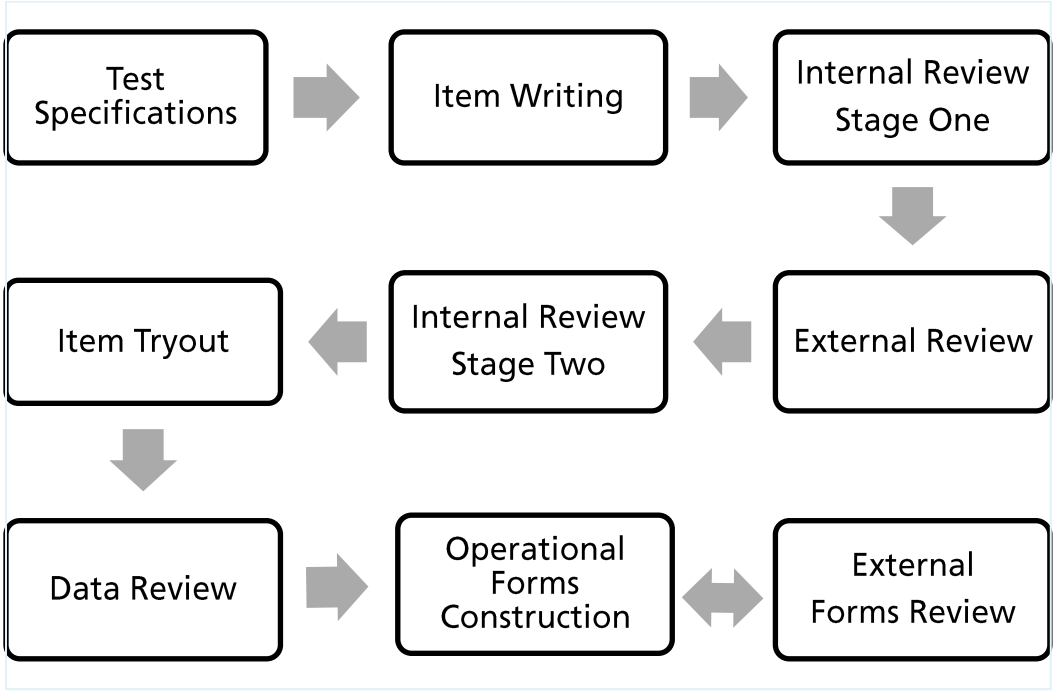
All types of validity evidence reflect on the quality of the test, but they do not guarantee its validity. They do not prove that the test measures what it purports to measure. They certainly cannot reveal whether the things being measured are those that ought to be measured. A high reliability coefficient, for example, shows that the test is measuring something consistently but does not indicate what that "something" is. For example, one can build a highly reliable mathematics test by including only simple computation items, but this would not be a valid test of problem-solving skills. Similarly, a poor test may show the same distribution of item difficulties as a good test, or it may show a higher average index of discrimination than a more valid test. The content specifications and the actual distribution of content coverage to the examinee need to be understood in order to clarify/disentangle the relationship between validity and reliability.

Carefully developed large-scale assessments should be supported by a continuous program of research and evaluation. Rational judgment also plays a key part in evaluating the validity of achievement tests against content and process standards and in interpreting statistical evidence from validity studies.

Test Development Procedures

The *Next Generation Iowa Assessments* are the result of an extended, iterative process during which draft test materials are developed and administered to national and state samples to evaluate their measurement quality and appropriateness. All materials for the *NGIA* have been part of an extensive field testing program throughout the state of Iowa. Figure 1 shows the process involved in test development.

Figure 1. Steps in Development of the *Iowa Assessments*



Test Specifications

Test specifications are created that outline (among other attributes) the statistical specifications for a test; the distribution of content, skills, item formats and cognitive levels across the test form; the test's organization; and any special accommodations and other conditions of test administration. Establishing these parameters beforehand helps to ensure the new forms are comparable to existing forms to the degree desired.

For the *NGIA*, comparability and continuity with the state proficiency expectations and college or career readiness benchmarks are part of the plan for both development and psychometric research. As test development proceeds, the test specifications are continually revisited and evaluated in an iterative process with Iowa educator involvement to ensure that the materials available for assembly of final forms reflect the evolving purposes of the assessments.

Item Writing

Items and stimulus/item sets (reading passages, graphs, maps, tables, and so on that support a group of items) are created according to the test specifications. Content specialists at ITP convene item-writing workshops and train educators on sound item writing practices. Educators are assigned to write items in the content areas and grade levels that best align with their experience in the classroom. Item production goals ensure a significant surplus of items across subject areas at each cognitive level so that the pool of available items in each subject and at each level is far greater than is needed to build each test. This surplus allows content experts to retain those items which exhibit the characteristics desired.

Internal Review—Stage One

After items are written, content specialists review them for content accuracy, fairness, and Universal Design considerations. The goal of these reviews is to make sure the items are accurate, fair, and accessible to all student subgroups in the diverse population of test takers. The items and associated materials are edited to ensure that they are clearly written and that reading loads are grade appropriate. The items are also copyedited for grammar and spelling at this stage in the process.

External Review

Once the items have been reviewed internally, ITP convenes panels of Iowa educators to review the items and associated stimuli (reading passages, tables, graphs, maps, and so on). After a formal training session in the review process, educators review the items for grade level appropriateness, content relevance, and accuracy. Since they have not been involved in the development process up to this point, external reviewers provide an objective "cold read" of potential test materials. A main goal of the educator review is to confirm that the items are appropriate for the intended level and content area. Since the adoption of the Common Core in 2010, approximately 400 Iowa teachers have reviewed and aligned tryout items.

Internal Review—Stage Two

ITP development staff reviews the items again after the educator panel review. This review focuses on edits made to the items during previous steps in the process and again checks for content accuracy, fairness, and Universal Design considerations.

Employment of Principles of Universal Design

ITP's content specialists are experienced in employing the Principles of Universal Design when developing items and tests so that all students have accessibility to the assessments in an equitable manner. This approach better ensures that all students can participate in the assessments. According to the NCEO Synthesis Report 44, there are seven elements of universally designed assessments. These elements are:

- Inclusive assessment population
- Precisely defined constructs
- Accessible, nonbiased items
- Amenable to accommodations
- Simple, clear, and intuitive instructions and procedures
- Maximum readability and comprehensibility
- Maximum legibility

All items for the *NGIA* are being or will be developed while taking into account these elements. ITP ensures development of items in accordance with these principles in the following manner:

- Items are developed to include a wide array of contexts and cultures. By developing these types of items, ITP ensures the participation of students with diverse learning needs.

- The test and item specifications serve as a map for precisely defining the constructs that the tests will measure. These specifications indicate to the item writer, content reviewer, and test development specialist exactly what is to be measured. The item could be assessing a particular part of a standard or a combination of elements within a standard. The reviews serve as a method for eliminating items that include assessment of knowledge outside of the standard. For example, mathematics items will have nonmathematical vocabulary below grade level; otherwise the tests would also assess reading ability, decreasing the validity of the measure.
- The review of items will serve to ensure that all items are accessible. Teacher review committees will have an opportunity to review the instructions to ensure that they are appropriate for the grade levels and subject areas of the students.
- Finally, ITP has experience in developing text, art, tables, maps, and diagrams with maximum legibility. We ensure the recruitment and inclusion of expertise on how text, art, tables, maps, and diagrams may be more easily understood by students with special needs.

Item Tryout

ITP collects data on the performance of the items that have successfully passed the review process by conducting a field test to determine how well the items are likely to perform operationally. Iowa students complete the field tests when they take the operational tests in numbers sufficient to ensure the associated statistical results are sound. Tryout of test materials and alignment to the core standards provide data necessary to ensure optimal placement of items for the measurement of growth. Since the adoption of the Common Core in 2010, approximately 432,000 Iowa students have responded to tryout items.

Data Review

The data collected during the tryout are analyzed for technical qualities related to item difficulty and discrimination. This analysis determines whether the items are appropriate measures of students' knowledge and the extent to which they will contribute to the test's overall reliability. Other aspects of the data review include key checks and the analysis of distractor choices, subgroup differences, and correlations with operational test forms. Only items that display acceptable descriptive statistics are eligible to appear on operational forms.

Operational Forms Construction

Items that survive data review become part of the pool of items that are eligible for use on operational tests. Procedures for creating the test forms are designed to ensure adequate content coverage in each subject area while being meaningful to students of varying achievement levels, as well as diversity of the items with regards to skill alignment, cognitive level, and difficulty. Careful attention is paid to item selection so that the final tests follow the predetermined test specifications and meet psychometric targets for difficulty, discrimination, and reliability.

Forms Review

Once draft tests have been constructed, they undergo another round of external reviews. Educators are recruited to evaluate the materials from a variety of perspectives, including appropriateness for the intended audience. A second group of experts evaluates materials for perceived fairness and sensitivity concerns. Members of this second group are selected from various ethnic and racial groups, both genders, and from among those who have expertise in the needs of student subgroups such as English language learners, students with special needs, and students who are visually impaired.

Predictive Validity and College Readiness

Tests such as the *Next Generation Iowa Assessments* are often used to support judgments about how well students are prepared for future instruction—that is, as general measures of readiness. Over the years, Iowa Testing Programs has conducted numerous studies to establish the predictive "power" of such assessments with respect to a variety of criterion measures, including high school GPA, college GPA, and scores on college entrance exams such as the ACT® and SAT® (e.g., Loyd, Forsyth, and Hoover, 1980, Ansley & Forsyth, 1983; Iowa Testing Programs, 1999). The *Guide for Research and Development, Forms A and B* includes the details of these studies.

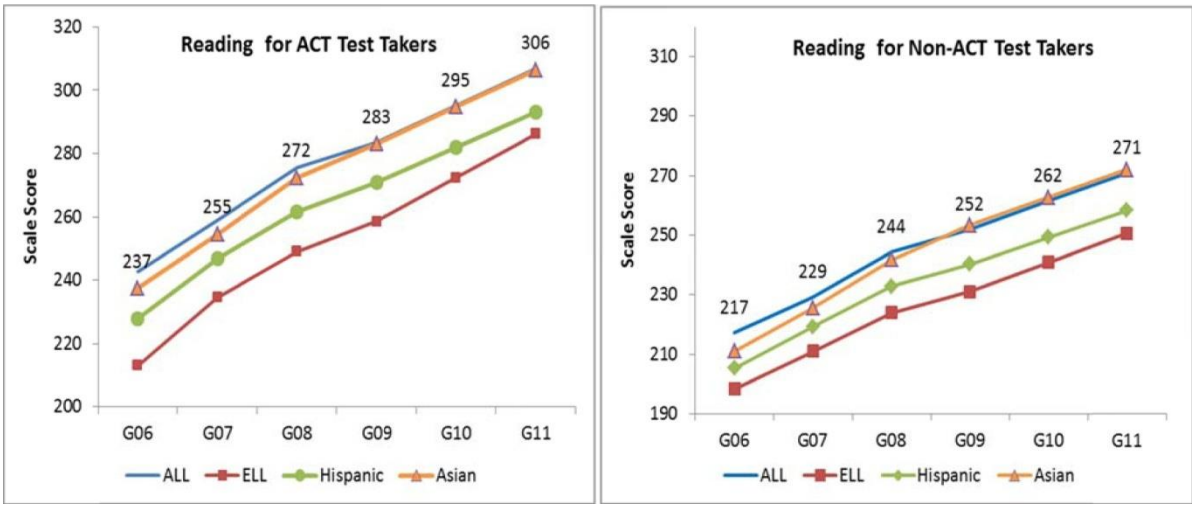
More recently, Furgol, Fina, and Welch (2011) investigated the relationship between performance on the pool of items being readied for the *NGIA* and college admissions test scores in a matched longitudinal cohort of over 25,000 students in grades 5 through 11 who tested annually over a five-year period. Evidence of a strong relationship between *Iowa Assessments* scores and the ACT composite score suggests that the *Iowa Assessments* and college readiness measures assess very similar if not the same achievement domains. This relationship sustains itself and strengthens from grades 5 to grade 11. Furgol, Fina, and Welch also reported the correlations between ACT and Iowa subject-area test scores for approximately 18,000 students in grades 8–11. These correlations are generally high across all grade levels and content areas providing supporting evidence for the use of the grade 11 scores to predict whether students are likely to meet or exceed the ACT College-Readiness Benchmarks. Note that the correlations between the grade 11 Iowa subject area tests and the corresponding ACT tests are as high as or higher than those between corresponding subject area tests on *EXPLORE*® and *ACT*.

This research will continue when the *NGIA* are introduced, and the college readiness standards will be statistically linked to the *NGIA*. In addition, as Iowa students complete their high school education and enroll in public (two-year and four-year) and private institutions in the state, ITP will validate the predictive validity of the *Next Generation Iowa Assessments* by tracking student performance throughout their postsecondary experiences.

In addition to providing evidence of content readiness or preparedness, this line of research is intended to provide additional validation of existing measures used in the state for admission to Regents institutions (for example, the RAI) and community colleges.

A study by Wang, Chen, and Welch (2011) examined group differences in the empirical trajectories of performance and established that growth trends for culturally (e.g. Asian and Hispanic) and linguistically diverse (i.e. English Language Learners) test takers run parallel to the college readiness trajectories identified by Furgol, et al. (2011). All effect sizes for departure from parallel trajectories were extremely small, as suggested by the results shown in Figure 2. Such results provide evidence of the appropriateness of using the *NGIA* scale to track the college readiness of all students, including those belonging to subgroups.

Figure 2. College Readiness in Mathematics



Interpretation and Utility of Readiness Information

College readiness information can help educators and families determine whether students are on track to successfully complete first-year college coursework upon graduation from high school or whether additional remedial coursework and preparation may be necessary. It allows families and educators to monitor student progress from middle school through high school and allows flexibility to determine the appropriate improvement and support strategies for students as they plan for post-secondary education opportunities. Monitoring the use of readiness information of the type described here is an important responsibility at the local level. This information should be used in ways that inform instruction and enhance learning for students as they prepare for postsecondary education opportunities. In addition, it can be used to complement existing admission criteria and processes already used by the Iowa's higher education institutions. An example of this type of information is provided in Appendix C. ITP routinely prepares reports and utility information to assist in the interpretation of assessment results.

Validity in the Assessment of Growth

Score interpretations that provide for the assessment of student growth over time are an important aspect of large-scale assessment in education. The measurement of growth through the *Iowa Assessments* is based on the Iowa Growth Model and the underlying vertical scale used in reporting, the national standard score (NSS) scale. Vertical scaling is the term used for the process of linking assessments to describe student growth over time. Although methods of

vertical scaling can be complex, the goal is quite simple: to create a framework and metric for reporting the educational development of individuals and groups. Today, vertical scaling is needed for assessments of growth toward college and career readiness standards and for adaptive testing. In these applications, comparative information about results from assessments of different levels of difficulty is needed to build a vertical scale, and this is incorporated into the forms assembly approach for the *NGIA* previously described. Assembling test forms with an **evidence-based** approach to growth on established content standards is a key element in vertical scaling.

Assessments matched to content that are not vertically aligned across grade or that reflect an overly granular approach to domain definitions and content specifications may show irregular patterns of growth across grades for both individuals and groups.

Validity Framework and Statistical Foundation of Growth Metrics

Assessing a student's growth on a learning continuum requires measures aligned to broad content standards and a level of cognitive complexity appropriate for that child's stage of development. Developmental appropriateness is (1) guided by research and practice in the achievement domain (e.g. the major domains of the Common Core State Standards in English Language Arts), and (2) established through extensive field testing of assessment materials. Valid and reliable measurement of growth requires both.

The *Next Generation Iowa Assessment* will be based on the vertical growth scale of the *Iowa Assessments*. This will allow educators to use a vertical scale to track student growth and continuity toward readiness for post-secondary plans. The development of the vertical scale is guided by consideration of content, learning progressions and empirical connections between scale points and college readiness benchmarks.

The vertical scales that have been constructed include a number of defining technical characteristics (Patz, 2007), including:

- An increase in difficulty of associated assessments across grades,
- An increase in scale score means with grade level, and
- A pattern of increase that is regular and not erratic.

The scale was developed using standard scores that describe a student's location on an achievement continuum, much like a learning progression for a broadly defined content domain. Expectations for a student's annual growth (beginning at any point on the scale) can be established based on intervention and instructional strategies. The scale tracks year-to-year growth and compares student expectations to achieved growth. The score scale is a vertical scale that quantifies and describes student growth over time. The current vertical scale, developed by the Iowa Testing Programs, is psychometrically sound, has been used extensively at the district and state level, and meets the technical requirements of large scale assessment (APA, AERA, NCME Standards, 1999).

Costs to School Districts and the State for Providing and Administering

The total program costs to school districts and the State of Iowa are yet to be determined for either the *Next Generation Iowa Assessments* (*NGIA*) or the Smarter Balanced Assessment Consortium (SBAC). As a frame of reference for the Assessment Task Force, Table 9 outlines the cost information available to date concerning these two options. However, please note that as states seek solutions to the delivery and implementation of SBAC, costs may change dramatically from the current estimates. Based on Iowa Testing Programs' (ITP)'s experience in the delivery, administration, scoring and reporting of assessments, the reasonableness of the $22.50 figure is questionable. States have also begun the process of securing bids through procurement processes to operationalize the delivery of the SBAC content. For example the state of Missouri recently awarded a contract for the delivery of SBAC content in ELA and mathematics. Missouri's per student costs included $6.20 to CRESST for ongoing development and $33.43 to a testing vendor for the online delivery, scoring and reporting of the SBAC tests. The total of $39.63 per student per grade per year was considerably higher than the estimates provided by SBAC.

As suggested in Table 9, ITP would continue to provide all item and test development, research and psychometrics associated with the *NGIA* to the State and districts within the state at no cost. For the administration, scoring and reporting services, only the direct costs to operate the program would be incurred by the State or districts to deliver and score the *NGIA*. In the current paper-based model, districts are charged approximately $4.00 per student for the packaging, shipping, receiving and warehousing of test materials; and the processing, scoring and reporting of test results. No additional overhead or other expenses are included in or added to the costs associated with these services.

The *NGIA* administration, scoring and reporting services can be configured to best match the State's need for information, professional development opportunities and costs. Table 10 outlines four different scoring and reporting models that vary in the purposes that they support, as well as the amount of information that is provided. Each model assumes a different approach to scoring, from central scoring to a combination of local and central scoring. Assumptions are also made with respect to the use of automated scoring engines for a number of the constructed-response items.

Model 1 assumes that the assessment is administered in a paper/pencil mode and all constructed-response and extended-response items would require human readers. Model 2 assumes that all student responses to the multiple-choice and technology-enhanced items would be scored by a centralized scoring system. Responses to the constructed-response items would be scored by human readers when automated scoring engines are not a viable alternative. Responses to the constructed responses items in ELA and Mathematics could be scored by a single reader with a percent of the responses being scoring by a second independent reader for quality checks or they could be scored by two independent readers, depending upon the cost implications. Given the complexity of the responses to the extended-response items, ITP is recommending that these be scored by two independent readers.

Models 3 and 4 offer two additional options (paper/pencil and online) that assume a combination of central and local scoring, which would enhance the opportunities for professional development and allow frequent release of the constructed-response items to provide the opportunity for students to receive additional diagnostic feedback on their responses.

As a next step in providing useful information to the Assessment Task Force concerning costs, ITP would encourage a discussion of the benefits and limitations of these four models, and the value of the information produced by each compared to the associated relative costs.

Perspective on Reporting Results

The ultimate goal of any testing program, large or small, is to provide results that assist educators in making informed decisions about improving curriculum, guiding instruction, and enhancing educational opportunities for their students. Accountability is predicated on this principle. Best practice in assessment, then, should place high priority on transmitting accurate, useful information. Reports of results are the vehicle for delivering that information.

Iowa Testing Programs strongly believes that reporting needs to provide timely and useful information to a variety of audiences including teachers, students, parents, administrators and policy-makers. Given the importance of scoring and reporting, the approach to both should be carefully considered and evaluated. Empowering teachers with the appropriate skills to better understand student evaluation and scoring is critical to impact on instruction and improvement.

ITP's many years of work in large-scale assessment have focused on ensuring that student testing data provide meaningful feedback to teachers, students and parents. Assessment data for these audiences must meet three principal requirements:

1. The reported results must be accurate.
2. The reports must be designed to complement classroom instruction.
3. The results must be presented in a format that can be easily interpreted by classroom teachers.

The technical systems assurance plan for reporting that we have designed to meet these three fundamental criteria is based on decades of experience with large-scale assessment reports. The ITP reports for the *NGIA* will be attractive, clear, easy to read, and easy to understand.

The scoring models agreed upon between the State of Iowa and ITP during implementation will determine the operational costs to deliver and score these assessments. Table 10 provides an overview of the models of administration and scoring available for consideration by the State with the *NGIA*. Score reports will be available in a web-based system and as paper-based reports with the information they contain dependent on the model selected. Web-based reports will be provided so that educators are able to access, save, and print them as needed. It should be noted that features of the four models described can be combined as desired.

Table 9. Comparison of Available Costing Information

Source	Development of Original Content	Development of Ongoing Content	Administration	Scoring and Reporting
Next Generation Iowa Assessments	No Cost	No Cost	Model Dependent	Model Dependent see Table 10
Smarter Balanced Assessment Consortium Assessments	Federally funded Race to the Top (RTTT) award to consortium	$6.20/student/year through contract with UCLA's National Center for Research on Evaluation, Standards, and Student Testing (CRESST) to maintain item and test pool	Unavailable at this time; bids for these costs received by other states are available. SBAC has posted at their website the following text: "Smarter Balanced has released cost estimates for its assessments that include expenses for ongoing research and development of the assessment system, as well as test administration and scoring. The end-of-year summative assessment alone is estimated to cost $22.50 per student. These costs are less than the amount that two-thirds of the Consortium's member states currently pay. **These costs are estimates because a sizable portion of the cost is for test administration and scoring services that will not be provided by Smarter Balanced; states will either provide these services directly or procure them from vendors in the private sector."** **http://www.smarterbalanced.org/resources-events/faqs/**	

Table 10. Comparison of Four Different Scoring and Reporting Models for *NGIA*

	Model 1	Model 2	Model 3	Model 4
Primary Purpose	Accountability	Accountability	Accountability and Informing Instruction	Accountability and Informing Instruction
Administration mode	Paper/pencil	Online	Paper/pencil	Online
Testing items	Secure	Secure	Partial release	Partial release
Scoring	Centrally scored Trained readers	Automated scoring engine (ASE) for all items that can be validly accommodated by this approach Centrally scored with trained readers for those that cannot be validly scored by ASE	Combination of central and local scoring (building, district, AEA) Professional development opportunities for teachers Scoring materials provided to teachers	Automated scoring engine (ASE) for all items that can be validly accommodated by this approach Combination of central and local scoring (building, district, AEA) Professional development opportunities for teachers Scoring materials provided to teachers
Types of Report	Reports provided by ITP	Immediate online reports for ASE items; reports provided by ITP for all other items	Reports provided by ITP	Immediate online reports for ASE items, reports provided by ITP for all other items
Information Provided	All scores Students do not receive papers	All scores Students do not retain responses	All scores Student receives paper with diagnostic feedback	All scores Student receives paper with diagnostic feedback
Turnaround time	2 to 3 weeks	2 to 3 weeks	As determined by district	As determined by district

Technical Support to Implement the Assessment

The *Next Generation Iowa Assessments* (*NGIA*) can be administered in either a paper-based or online mode, with the decision residing at the district level. This approach gives schools flexibility in their use of technology resources and scheduling so that they can administer tests in ways that meet their needs.

The online system is web-based and will provide all needed security features as well as full support for the test-taking experience. Each student will have a logon and password to the tests within the online testing system. Teachers will facilitate the administration just as they do with a paper-based administration. Once the student completes the test, the responses are submitted via the online system for scoring.

One key advantage of a web-based system is that it does not require the installation of extensive software locally. It supports current releases of most major operation systems and browsers. The list of supported versions will continue to evolve between now and the introduction of this assessment in 2015. The table below identifies the system requirements as currently configured.

All devices must have the administrative tools and capabilities to temporarily disable features, functionalities, and applications that could present a security risk and question the integrity of the results. These include, but are not limited to, such functions as unrestricted Internet access, cameras, email and instant messaging, and screen captures.

Table 11. Typical System Requirements for Online Delivery

	Typical System Requirements
Processor	2 GHz or faster
RAM	2 GB RAM
Operating System	• Windows XP (SP3), Vista, Windows 7 and 8 • Mac OS 10.5 or higher
Web Browser	• IE 8 and higher • Firefox 17 and higher • Chrome (ESR 17 and later) • Safari 5.0 or higher
Tablets	• iPad with iOS5 or higher • Android 4.2 or higher
Minimum Screen Resolution	1024 × 768 (computer), scaling to a minimum of 10-inch (tablet)

Online Experience

Iowa Testing Programs has invited a number of schools (public and private) to participate in an initial online pilot administration in November, 2013. All tests were available online for all grades. The pilot continues as of this writing (see Appendix B for technical support documentation related to this pilot testing). Prior to this pilot, ITP began offering assessments via computer in Spring, 2011, with the introduction of the *Iowa End-of-Course* (*IEOC*) tests online. Over 70,000 students in Iowa (primarily high-school aged, but some middle-school students as well) have successfully taken tests online in the following areas: Algebra I, Algebra II, Geometry, Matrix Algebra, Probability and Statistics; Physical Science, Biology, Chemistry; English Language Arts; U.S. Government, and U.S. History. Beginning in 2012, ITP began offering the *Iowa Algebra Readiness Assessment* (*IARA*) online. This screener is predominantly used in the middle school grades to assist in the identification students who demonstrate readiness for Algebra. In a little over a year, more than 6,500 students in Iowa have been tested online with *IARA*. The results, available in real-time, have been invaluable in allowing schools to place the appropriate students into Algebra courses.

Paper/Pencil Administration

Students taking the paper-based administration of the *NGIA* will see the items in a test booklet and will respond to all items in a separate scannable answer booklet. The answer booklet will capture the constructed-response items in addition to the multiple-choice items.

ITP has substantial experience delivering, scoring and reporting results from a paper-based administration. ITP will draw from this experience for the paper-based administration of the *NGIA*.

Comparability of Paper-Based and Online Modes of Administration

In the spring of 2012 and the spring of 2013, comparability studies with the *NGIA* item pool comparing paper versus an online administration were conducted. The data were collected through a counterbalanced design that required that the same students take both formats of the items but in different orders to balance any practice effect. Students were randomly assigned (1) paper-and-pencil version first and computerized version second or (2) the computerized version first, paper-and-pencil version second. Approximately 12,000 student responses were obtained for each item in the pool in each administration format.

The study sought to determine the equivalence in both construct and scores (average student performance) across formats. In addition, differences in visual presentation, item position, and complexity of graphics were compared. Analysis included confirmatory factor analysis, differential item functioning statistics, item-level p-value and biserial analyses, and overall differences in performance.

The differences in performance between the two modes was minimal. No evidence was found to suggest any systematic advantage of taking the items in one mode over the other.

Accommodations

In the state of Iowa, IEPs or 504 Plans generally take precedence and dictate the conditions under which the *NGIA* are to be given. The following table identifies some of the available accommodations by administration mode.

Table 12. Accommodation Supports for Students

Area of Support	Available in Online Administrations	Available in Paper/Pencil Administrations
Presentation of Testing Materials	• Audio read-alouds • Magnification • Line Readers	• Read-alouds • Large Print • Braille
Student Interaction with Testing Materials	• Highlighting • Scratch Paper • Online Calculator or Personal Calculator	• Highlighting • Scratch Paper • Personal Calculator • Student dictates responses to transcriber
Testing Conditions	• Extra time • Flexible configuration of test order, breaks, environment • Individual or group administration	• Extra time • Flexible configuration of test order, breaks, environment • Individual or group administration

Reporting

Iowa's new assessment program is required to provide, "valid, reliable, and fair measures of student progress toward college or career readiness," (HF215, Sec. 47 b(2)), and this requires scoring and reporting categories that are meaningful with respect to postsecondary planning for a variety of audiences. Table 13 provides an overview of the approaches to scoring for the *Next Generation Iowa Assessments* and the SBAC assessments.

The *NGIA* are designed to support aggregate scores in ELA and Mathematics for accountability based on proficiency levels, however, that is only the starting point for the information available. Diagnostic information is provided through scores based on specific "claims" about what student know and are able to do, as well as domain scores that reflect student performance on the major strands of the Iowa Core. The domain scores are useful for teachers and schools in evaluating instruction relative to the Iowa Core. Such scores cannot be reported for the SBAC assessments in "SBAC certified" programs due to the potential for inadequate domain sampling in the operation of the adaptive algorithm for item selection. Other scores for growth and college or career readiness are based on continuing ITP research in the measurement of growth and preparation for the content of credit-bearing college courses and for postsecondary training in career-oriented programs. Finally, metrics familiar to parents and policymakers that reference student performance relative to national benchmarks are available for comparative purposes.

Scoring and reporting are integral aspects of the validity of information deriving from an assessment, and the approach for the *NGIA* is to offer multiple levels of reporting to reflect the likely multiple uses of assessment information. For a summative, accountability assessment, scoring and reporting for proficiency against the ELA and Mathematics domains of the Iowa Core is critical, as is comparability of proficiency-based inferences within-grade and between-grades *for all students*. The use of fixed-form assessments is the most robust, transparent, and sustainable design approach to carry the argument for validity of proficiency-based inferences over time. Adaptive testing designs require extremely large item pools within and between grades to ensure adequate content coverage of the standards *for all students*, and they rely on complex algorithms and item selection rules that make validating the alignment of any particular student's actual test to the Iowa Core virtually impossible. In addition, information that could inform instruction consistently for all students may be difficult to obtain with an adaptive approach because individuals are exposed to so many different combinations of items.

Table 13. Reporting Metrics for the *NGIA* and SBAC Assessments

	ELA	Mathematics	Other Domains
Next Generation Iowa Assessments	Claim Scores Domain Scores Proficiency Levels Growth Indicators College/Career Readiness National Comparisons	Claim Scores Domain Scores Proficiency Levels Growth Indicators College/Career Readiness National Comparisons	Social Studies Science
Smarter Balanced Assessment Consortium Assessments	Claim Scores Proficiency Levels Comparison to SBAC states	Claim Scores Proficiency Levels Comparison to SBAC states	

The *NGIA* will report on a variety of metrics to assist various audiences in understanding and using this information. The following tables show the reporting content categories, by grade, for each test that will define the Domain Scores listed in Table 13.

Table 14. *NGIA* ELA Reporting Domain Categories

Test	Reporting Categories	2	3	4	5	6	7	8	HS
ELA – Part 1, Reading	Total Reading Score	✓	✓	✓	✓	✓	✓	✓	✓
	Key ideas	✓	✓	✓	✓	✓	✓	✓	✓
	Craft and Structure	✓	✓	✓	✓	✓	✓	✓	✓
	Integration of Knowledge and Language	✓	✓	✓	✓	✓	✓	✓	✓
ELA – Part 2, Language	Total Language Score	✓	✓	✓	✓	✓	✓	✓	✓
	Conventions of Spoken English	✓	✓	✓	✓	✓	✓	✓	✓
	Knowledge of Language	✓	✓	✓	✓	✓	✓	✓	✓
	Vocabulary Acquisition and Use	✓	✓	✓	✓	✓	✓	✓	✓
ELA – Part 3, Writing	Total Writing Score	✓	✓	✓	✓	✓	✓	✓	✓
ELA Composite	Total Score	✓	✓	✓	✓	✓	✓	✓	✓

Table 15. *NGIA* Math Reporting Domain Categories, Grades 3 – 8

Reporting Categories	3	4	5	6	7	8
Operations and Algebraic Thinking	✓	✓	✓			
Number and Operations in Base Ten	✓	✓	✓			
Number and Operations -- Fractions	✓	✓	✓			
Measurement and Data	✓	✓	✓			
Geometry	✓	✓	✓	✓	✓	✓
Ratios and Proportional Relationships				✓	✓	
The Number System				✓	✓	✓
Expressions and Equations				✓	✓	✓
Statistics and Probability				✓	✓	✓
Functions						✓
Mathematics Problem Solving and Data Interpretation	✓	✓	✓	✓	✓	✓
Mathematics Concepts	✓	✓	✓	✓	✓	✓

Scoring Quality Assurance

ITP's commitment to producing high-quality products and services is reflected in our ongoing efforts to implement methods that ensure not only the accuracy of our test items, but also the individual and summary data that make up our score reports. Individual student and summary data are examined against a variety of sources as well as cross-referenced to ensure completeness and accuracy. We have instituted a multi-step scoring system that stresses quality, efficiency, flexibility, and security. Documentation and logs detailing the series of checks conducted throughout the data operations process are compiled and can be provided to school districts.

Before the release of assessment results, the accuracy and completeness of the data that make up the score reports are scrutinized. The quality checks conducted on the final data encompass not only the final output of score reports but also the production scoring system:

- Answer keys: Frequency distributions are produced for actual student response data to verify the accuracy of the multiple-choice answer keys.
- Scoring and reporting system: An independent statistical program is used to verify the scoring process.
- Data used to produce the final reports: All data elements are checked to ensure that no data are missing, that processing and scoring rules were applied correctly, and that the file layout is correct.

References

American Educational Research Association, American Psychological Association, & National Council of Measurement in Education. (1999). Standards for Educational and Psychological Testing. Washington, DC: Author.

Ansley, T. N. and Forsyth, R.A. (1983). Relationship of elementary and secondary school achievement test scores to college performance. *Educational and Psychological Measurement,* 43: 1103–1112.

Common Core State Standards Initiative. (2012). Common Core State Standards for English Language Arts. Retrieved from http://www.corestandards.org/the-standard

Common Core State Standards Initiative. (2012). Common Core State Standards for Mathematics. Retrieved from http://www.corestandards.org/the-standard

Furgol, K., Fina, A., & Welch, C. (2011). Establishing *validity evidence to assess college readiness through a vertical scale*. Paper presented at the 2011 Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.

Linn, R. L., Baker, E. L., and Dunbar, S. B. (1991). Complex performance-based assessments: Expectations and validation criteria. *Educational Researcher,* 20: 15–21.

Loyd, B. H., Forsyth, R. A., and Hoover, H. D. (1980). Relationship of elementary and secondary school achievement test scores to later academic success. *Educational and Psychological Measurement,* 40: 1117–1124.

Patz, R. J. 2007. *Vertical Scaling in Standards-Based Educational Assessment and Accountability Systems*. Prepared for the Technical Issues in Large-Scale Assessment (TILSA) State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO).

Appendix A: Excerpts from Commissioned Alignment Study

From page 5 of the report (ITP emphasis added):

"The alignment studies for the Iowa Assessments for grades 3-8, 10 and 11 mathematics and reading were held on March 27-28, 2013, in Des Moines, Iowa. The purpose of each alignment study was to determine the degree of alignment between the Iowa Core Standards for each grade and the test items found on the corresponding grade-level Iowa Assessment. **It should be noted that the Iowa Core Standards used for the alignment process were not the standards used to develop the Iowa Assessments**. The Iowa Department of Education provided the Iowa Core Standards to be used for the study. Each alignment study involved a group of independent third-party reviewers whose role was to judge the depth-of-knowledge level of each standard, reach consensus, and then-independently judge the depth-of-knowledge level of each test item, including identifying the primary and possibly a secondary standard to which each item was aligned."

From page 15 of the report (ITP emphasis added):

Interpretation of Mathematics Alignment Results

Reviewers were asked to align the assessment items to the Iowa Common Core. As previously mentioned in the Overview**, the standards of the Iowa Common Core were not the standards used to develop the test but were used in this study as requested by the Iowa Department of Education. Using standards that were not used to develop the items could affect the results of the alignment criteria.** There could be variability in assigning the same standard to an item by the reviewers. This could be a result of asking the reviewers to align items to standards that the assessment was not originally intended to measure and to continue the process until the reviewers found an alignment in one or more grades below the original grade.

Appendix B: Technical Documents to Support the 2013 Online Pilot Administration

Iowa Assessments Online Test Administration
Technical Setup and System Requirements

Review your firewall or proxy server settings

- IP addresses are set as approved/unblocked or given the highest priority.
- URLs have been approved and given unrestricted access.

IP Address and Port Listing

***DataManager* Component**	**URL**	**External IP**	**Port**
DataManager	http://www.riversidedatamanager.com https://www.riversidedatamanager.com	209.235.124.49	80 443
DataManager Online Testing	https://www.riversideonlinetest.com http://starttest.com https://starttest.com http://starttest2.com https://starttest2.com NOTE: If you experience problems with the URLs above, try entering them as follows: http://starttest.com http://starttest2.com *.starttest.com* *.starttest2.com*	209.235.124.50 64.27.100.27 64.27.64.232 66.70.68.224 206.188.17.0/24 64.106.193.0/24 NOTE: The last two addresses in the list above are IP ranges. If your system does not accept '0/24', then try entering the address without the /24.	80 443
System Check	http://www.riversideonlinetest.com/systemcheck http://dmsystemcheck.programworkshop.com		80 443

Online Testing Student Workstation System Requirements

Student System Requirements for Online Testing			
Operating System	**Web Browser**	**Software**	**Screen Resolution**
Windows 7 Windows Vista Windows XP	Microsoft Internet Explorer 7 and up (IE 8 and up recommended) JavaScript must be enabled	Adobe Reader 9.0 or higher	Minimum: 1024x768 screen resolution Should not exceed 1440x900. Windows: DPI Scaling to default (96 dpi) Mac: DPI Scaling to default (72 dpi)
Mac OS X 10.5 Leopard or higher	Safari 5 JavaScript must be enabled	Adobe Reader 9.0 or higher	
-PC: 1 GHz or faster processor -PC and Mac: 1 GB RAM or greater			

Check your bandwidth and System Requirements compliance

Run the System Check utility to ensure computers in each lab meet the system requirements. The utility will also indicate the number of simultaneous users the school's network can support. When you run the utility, you can ignore the warning about pop-ups; pop-ups are not used for test administration.

You can access the System Check utility at: http://www.riversideonlinetest.com/systemcheck

Download the Online Testing Secure Browser

The Secure Browser opens the student login page using a browser window (Internet Explorer or Safari). It prevents the student from using other programs on the workstation while the testing browser is open. This preserves the integrity of the test and prevents students from accessing other websites and performing certain actions that could interrupt the test or cause invalid test results.

The Secure Browser is a small, stand-alone executable file that can run from the desktop. It does not require an install. The executable file needs to be placed on each student workstation that will be used for testing.

You can access the secure browser download here:
For Windows: Secure Student Browser - Windows
For Macs: Secure Student Browser - Mac

1. Click the appropriate link
2. Open the .zip file (you can also do a "Save as" to save the .zip file in a more convenient location for retrieval to download to each computer)
3. Drag the .exe file onto the computer's desktop

DataManager System Requirements (Teacher/Proctor interface)

<table>
<tr><th colspan="4">General DataManager System Requirements</th></tr>
<tr><td>Operating System</td><td>Web Browser</td><td>Software</td><td>Screen Resolution</td></tr>
<tr><td>Windows 7
Windows Vista
Windows XP</td><td>Microsoft Internet Explorer 7 and up (IE 8 and up recommended)
or Firefox 5 or higher
JavaScript must be enabled</td><td>Adobe Reader 9.0 or higher
Adobe Flash player for training videos</td><td rowspan="2">Minimum:
1024x768 screen resolution</td></tr>
<tr><td>Mac OS X 10.5 Leopard or higher</td><td>Safari 5
or Firefox 5 or higher
JavaScript must be enabled</td><td>Adobe Reader 9.0 or higher
Adobe Flash player for training videos</td></tr>
<tr><td colspan="4">System Requirements for reporting
- PC: 2 GHz or faster processor
- PC and Mac: 2 GB RAM or more is recommended for optimal performance</td></tr>
<tr><th colspan="4">Internet/Network Requirements</th></tr>
<tr><td colspan="4">High Speed Internet connection</td></tr>
</table>

Staff/User File Setup:

Guidelines and Layout for Submitting Staff Information

Iowa Assessments Online Pilot – Fall 2013

Guidelines for Submitting Files for the Iowa Assessments Online Pilot – Fall 2013

To submit student information for those who will take the Iowa Assessments online, and register staff that will proctor the online tests, you will need to upload 3 files to the Iowa Testing Program's Online Tools website – a student bar code file, a staff/user file, and a location file. This guide is designed to provide instruction on creating the Staff/User file that will be used to set up user accounts for staff that will proctor the tests. Below is a general overview of how to submit the files once they're ready, on the following pages are field requirements and descriptions for the Staff/User file.

The 3 files need to be in .xls, .xlsx, .csv or .txt file format.

To submit your files once they're ready, go to the Iowa Testing Programs website: https://itp.education.uiowa.edu/.

Click "Bar Coding" in the menu on the left. At this point the site will ask for your username and password to log into the secure Online Tools Bar Coding site.

Once you're logged in the first step is to create a new bar code order – click "Begin new bar code order".

This page has a series to tabs to enter details about this bar code order. Go through each tab and enter the requested information. *Note: under the "Test" tab you will see a drop-down box, open the menu and you should see the option "Online Pilot IA". Select "Online Pilot IA".

After you've clicked "Save" within the Save tab you should see a green "SUCCESSFUL" box with a link to go to the File Management Center, click that.

At this point you have successfully created the bar code order and the page you're on now is where you will upload the 3 files we need to prepare your account for online administration. Click Browse, locate and select/open the file to upload, click Upload. Repeat this process two more times to submit all 3 files to this bar code order.

Here are the headers for the Staff/User file template:

District / Area*	School/ Building	School/ Building Code	Class	Grade	Code	First Name *	Last Name *	Middle Name	Gender	E-Mail Address *	Password	Roles*	Active User*

Here are the specifications for the fields:

Field	Data Type	Required?	Max. Length/Format	Valid Values / Examples
District/Area*	Alphanumeric	Yes	30	A-Z, 0-9, Blanks, are acceptable
School/Building	Alphanumeric	No	30	A-Z, 0-9, Blanks are acceptable
School/Building Code	Alphanumeric	No	15	A-Z, 0-9, Blanks, dashes and period are acceptable. May be used to avoid duplicate building names.
Class	Alphanumeric	No	30	A-Z, 0-9, Blanks are acceptable. Class names must be unique to each class.
Grade	Numeric	No	2	00 to 13
Code	Alphanumeric	No	5	A-Z, 0-9, Blanks, dashes and period are acceptable. May be used to avoid duplicate building/class names.
First Name *	Alphanumeric	Yes	20	A-Z, 0-9, Blanks or dashes acceptable
Last Name *	Alphanumeric	Yes	20	
Middle Name	Alphanumeric	No	20	

Field	Data Type	Required?	Max. Length/Format	Valid Values / Examples
Gender	Alphanumeric	No	1	"M" or "F"
E-Mail Address *	Alphanumeric	Yes	50	Must follow valid email id format e.g., tomjones@abcschool.com
Password	Alphanumeric	No	8 - 10	Min 8 characters, max 10 characters, must contain one number.
Roles*	Alpha/Numeric	Yes	7	A, T, P, D are acceptable • A: Administrator • T: Teacher • P: Proctor • D: Digital Resource Access
Active User*	Alpha/Numeric	Yes	1	'Y' or 'N'; blank defaults to 'Y'

Location File Setup:

Guidelines and Layout for Submitting Location Information

Iowa Assessments Online Pilot – Fall 2013

Guidelines for Submitting Files for the Iowa Assessments Online Pilot – Fall 2013

To submit student information for those who will take the Iowa Assessments online, and register staff that will proctor the online tests, you will need to upload 3 files to the Iowa Testing Program's Online Tools website – a student bar code file, a staff/user file, and a location file. This guide is designed to provide instruction on creating the Location file that will be used to establish relationships between classes and school buildings within the school system. Below is a general overview of how to submit the files once they're ready, on the following pages are field requirements and descriptions for the Location file.

The 3 files need to be in .xls, .xlsx, .csv or .txt file format.

To submit your files once they're ready, go to the Iowa Testing Programs website: https://itp.education.uiowa.edu/.

Click "Bar Coding" in the menu on the left. At this point the site will ask for your username and password to log into the secure Online Tools Bar Coding site.

Once you're logged in the first step is to create a new bar code order – click "Begin new bar code order".

This page has a series to tabs to enter details about this bar code order. Go through each tab and enter the requested information. *Note: under the "Test" tab you will see a drop-down box, open the menu and you should see the option "Online Pilot IA". Select "Online Pilot IA".

After you've clicked "Save" within the Save tab you should see a green "SUCCESSFUL" box with a link to go to the File Management Center, click that.

At this point you have successfully created the bar code order and the page you're on now is where you will upload the 3 files we need to prepare your account for online administration. Click Browse, locate and select/open the file to upload, click Upload. Repeat this process two more times to submit all 3 files to this bar code order.

Here are the headers for the Location file template:

District/ Area	School/ Building	School/ Building Code	Class	Grade	Class Code	Address	City	State	Zip Code

Here are the specifications for the fields:

Field	Data Type	Required?	Max. Length	Valid Values / Examples
District / Area*	Alphanumeric	Yes	30	A-Z, 0-9, Blanks are acceptable
School/Building*	Alphanumeric	Yes	30	A-Z, 0-9, Blanks are acceptable
School/Building Code	Alphanumeric	No	15	A-Z, 0-9, Blanks, dashes and period are acceptable. May be used to avoid duplicate building names.
Class*	Alphanumeric	Yes	30	A-Z, 0-9, Blanks are acceptable. Class names must be unique to each class.
Grade*	Numeric	Yes	2	00 to 13
Code	Alphanumeric	No	5	A-Z, 0-9, Blanks, dashes and period are acceptable. May be used to avoid duplicate class names.
Address	Alphanumeric	No	40	A-Z, 0-9, Blanks or dashes acceptable
City	Alphanumeric	No	20	A-Z, 0-9, Blanks or dashes acceptable
State	Alpha	No	2	'IA' 'IL' or 'CA' etc.
Zip Code	Numeric	No	10	Five Numeric or Five Numeric plus '-' and Four numeric (e.g.12345-6789).

Appendix C: Example of Utility of Information Prepared by Iowa Testing Programs

The University of Iowa
ITP
Iowa Testing Programs

Iowa Assessments

How Ready are Iowa Students?

Preparing students to be college or career-ready is critically important in a highly competitive global economy. *One Unshakable Vision* (2011) identifies a need for such information for all students.

Recent research has focused on the readiness of students who are college bound. This report focuses on all students in the Iowa Class of 2012. Iowa Testing Programs provides this information to raise awareness of the importance of monitoring progress towards readiness of all students for post-secondary opportunities.

This report traces progress of the graduating class of 2012 from 6th grade through 11th grade. It answers the following important questions:

- In the Iowa class of 2012, what percent of students are ready for college when they graduate?
- In the class of 2012, what percent of students were on track for readiness as younger students (for example, as 6th graders, 7th graders or 8th graders)?
- In the class of 2012, what percent of Iowa students are proficient? Why is there a difference between proficiency and readiness?
- How can information about post-secondary readiness be used?

The University of Iowa
ITP
Iowa Testing Programs

Iowa Testing Programs
College of Education
University of Iowa
Iowa City IA 52242–1529

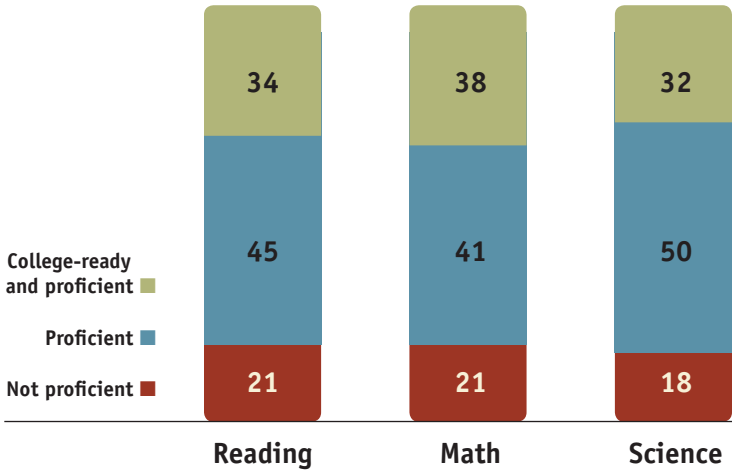
steve–dunbar@uiowa.edu
catherine–welch@uiowa.edu

www.education.uiowa.edu/itp

In the Iowa class of 2012, what percentage of students are ready for college when they graduate?

For **all** Iowa students graduating in the spring of 2012, Iowa Assessment data suggest that 34% are ready for college coursework in reading, 38% in mathematics and 32% in science. Contrast this with similar percentages for **college-bound** students where 62% are ready in reading, 52% in mathematics and 40% in science (ACT, 2010).

Figure 1: All Iowa Students



Some student groups within Iowa are more prepared than others. The table below provides the percent of Iowa students from the class of 2012 prepared for college-level courses.

	Reading	Math	Science
All students	34	38	32
African American	12	10	7
American Indian	12	17	12
Asian	36	46	37
Hawaiian/Pacific Islander	23	37	20
Hispanic	18	17	14
White	38	41	35
Two or more races	34	29	26
Male	32	44	33
Female	39	33	31
English Language Learners	2	7	3
Students eligible for free or reduced-pricedlunches	19	20	15

References

ACT. (2011). *The condition of college and career readiness Class of 2011 Iowa*.

Iowa Department of Education (2011). *One Unshakable Vision*.

ACT. (2008). *The Forgotten Middle*.

In the class of 2012, what percent of students were on track for readiness as 6th graders, 7th, graders or 8th graders?

Research would suggest that waiting for high school is too late to learn whether students are on track for college or a career. ACT researchers found that the "level of academic achievement that students attain by 8th grade has a larger impact on their college and career readiness by the time they graduate from high school than anything that happens academically in high school" (ACT, 2008).

Longitudinal data available on the Iowa Assessments™ allow the monitoring of the performance of this class between 6th grade and 12th grade. **Figures 2**, **3** and **4** capture the readiness of the class as 11th graders (as previously indicated in **Figure 1**), but also provide the percent of those students that were on track for readiness in earlier grades in reading, mathematics, and science. For example, **Figure 2** indicates that 33% of the Class of 2012 were on track to being college ready in mathematics as early as 6th grade (G6). The other figures suggest that 35% of the Class of 2012 were on track in reading in 6th grade while only 26% of the students were on track in science.

Figure 2: Percent of Class of 2012 in Three Performance Categories: Reading

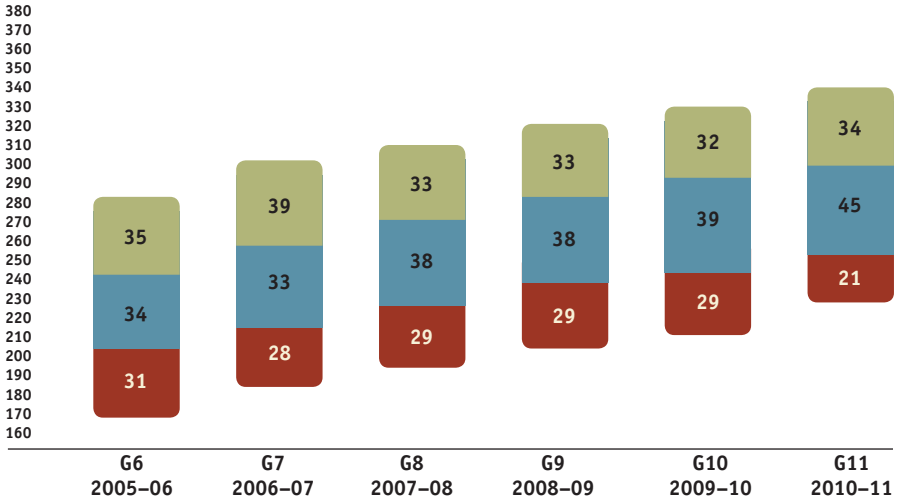
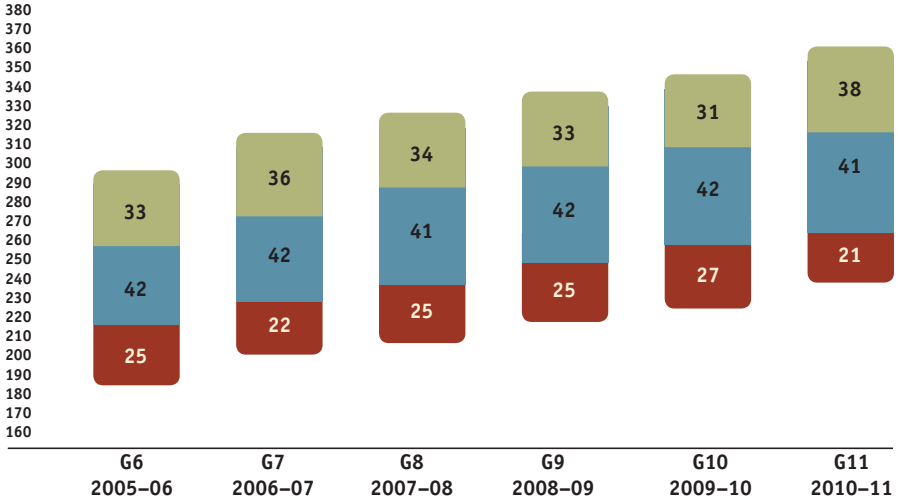


Figure 3: Percent of Class of 2012 in Three Performance Categories: Math

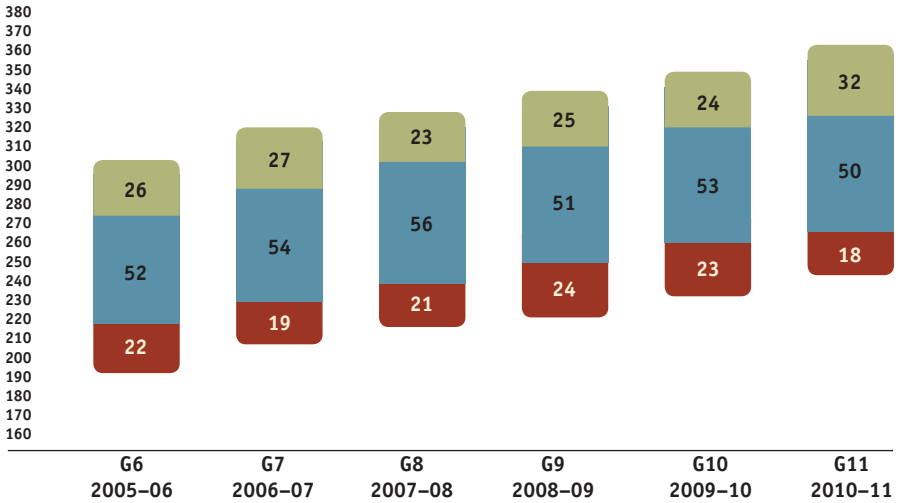


College-ready and proficient

Proficient

Not proficient

Figure 4: Percent of Class of 2012 in Three Performance Categories: Science



References

ACT. (2011). *The condition of college and career readiness Class of 2011 Iowa*.

Iowa Department of Education (2011). *One Unshakable Vision*.

ACT. (2008). *The Forgotten Middle*.

Why do proficiency and readiness indicators disagree?

In the class of 2012, what percent of Iowa students are proficient? **Figures 2**, **3** and **4** answer this question. Why is there a difference between proficiency and readiness? A critical element of The No Child Left Behind Act (2000) has been the annual identification of students who are proficient and those who are not proficient.

Students who are proficient have demonstrated an overall competence in fundamental knowledge and skills appropriate for their grade level as defined by the state's proficiency level descriptors (http://educateiowa.gov).

Figures 2, **3** and **4** also provide information about the percent of students from the class of 2012 who were defined as proficient and not proficient from 6th grade to 11th grade in reading, mathematics and science. For example, **Figure 2** illustrates that 25% of students were not proficient in mathematics in grade 6 and 21% remained not proficient by grade 11. Similar trends are found in reading and science.

These figures provide information to illustrate the contrast between the definitions of not proficient, proficient and college-ready. Proficient and college-ready standards introduce two different levels of preparation that are best illustrated by their definitions.

Definitions of Preparation for 6th Grade Students in Mathematics

College-ready and proficient	Students are on track in theirlevel of preparation to be successful in entry-level, credit-bearing courses in mathematics
Proficient, but not yet college-ready	Students can understand math concepts, solve word problems, and interpret data from graphs and tables. Students are sometimes able to use estimation methods.
Not proficient	Students can seldom understand math concepts, solve word problems, use estimation methods, or interpret data from graphs and tables.

References

ACT. (2011). *The condition of college and career readiness Class of 2011 Iowa*.

Iowa Department of Education (2011). *One Unshakable Vision*.

ACT. (2008). *The Forgotten Middle*.

How can this information be useful for future graduating classes?

Trends within the state for the past ten years suggest consistent results across graduating classes with respect to college readiness and proficiency. The results of this research suggest that future classes can use information from the Iowa Assessments™ to track growth towards readiness and predict performance on college entrance exams.

- **Policymakers** can use this information to compare Iowa performance to that of other states that test all students with the ACT
- **Policymakers** can use this information to gauge the readiness of future classes of students
- **Schools** can use this information to identify the appropriate places for interventions and change
- **Schools** can use this information to assess needs for coursework offerings
- **Schools** can use this information to evaluate effectiveness of programs and effects of program changes
- **Students, teachers, and parents** can use this information to set goals and monitor progress
- **Students, teachers, and parents** can use this information as they plan high school coursework
- **Students and teachers** can use this information to predict admissions decisions based, for example, on the Iowa Regents Admission Index
- **Researchers at Iowa Testing Programs** can expand this work to include a variety of programs at two- and four-year institutions.

References

ACT. (2011). *The condition of college and career readiness Class of 2011 Iowa*.

Iowa Department of Education (2011). *One Unshakable Vision*.

ACT. (2008). *The Forgotten Middle*.