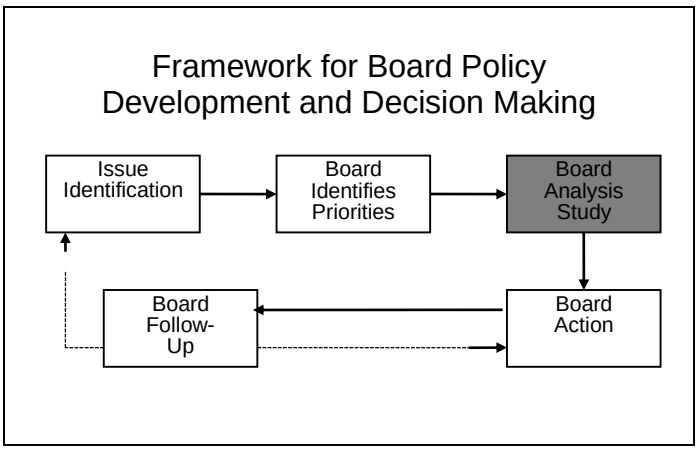
Iowa State Board of Education

Executive Summary

March 31, 2016



Agenda Item: Assessment Task Force Recommendation on Science Assessment

Iowa Goal: Goal 1: All children will enter school ready to learn.
Goal 2: All PK-12 students will achieve at a high level.

State Board Role/Authority Iowa Code section 256.7(21) requires the State Board to submit to the general assembly recommendations for modifications of assessments of student progress on core academic indicators. The State Board also has a priority related to assessment.

Presenter: Dave Tilly, Deputy Director
Division of Learning and Results

Attachments: 1

Recommendation: It is recommended that the State Board hear and discuss the recommendation from the Iowa Assessment Task Force regarding a Science Assessment for Iowa.

Background: This is a continuation of the conversation on the State Board's priority on assessment. Since the State Board adopted new Science Standards for Iowa, the assessment task force reconvened and deliberated on a recommendation for a new Science Assessment for Iowa. On March 11, 2016, the task force voted on a set of recommendations for a Science Assessment. These will be presented to the State Board for consideration.

2016 Assessment
Task Force Report - Science

March 21, 2016

State of Iowa
Department of Education
Grimes State Office Building
400 E. 14th Street
Des Moines, IA 50319-0146

State Board of Education

Charles C. Edwards, Jr., President, Des Moines
Michael L. Knedler, Vice President, Council Bluffs
Brooke Axiotis, Des Moines
Michael Bearden, Gladbrook
Diane Crookham-Johnson, Oskaloosa
Angela English, Dyersville
Rosie Hussey, Clear Lake
Mike May, Spirit Lake
Mary Ellen Miller, Wayne County
Hannah Rens, Student Member, Sioux City

Administration

Ryan M. Wise, Director and Executive Officer
of the State Board of Education

It is the policy of the Iowa Department of Education not to discriminate on the basis of race, creed, color, sexual orientation, gender identity, national origin, sex, disability, religion, age, political party affiliation, or actual or potential parental, family or marital status in its programs, activities, or employment practices as required by the Iowa Code sections 216.9 and 256.10(2), Titles VI and VII of the Civil Rights Act of 1964 (42 U.S.C. § 2000d and 2000e), the Equal Pay Act of 1973 (29 U.S.C. § 206, et seq.), Title IX (Educational Amendments, 20 U.S.C.§§ 1681 – 1688), Section 504 (Rehabilitation Act of 1973, 29 U.S.C. § 794), and the Americans with Disabilities Act (42 U.S.C. § 12101, et seq.). If you have questions or complaints related to compliance with this policy by the Iowa Department of Education, please contact the legal counsel for the Iowa Department of Education, Grimes State Office Building, 400 E. 14th Street, Des Moines, IA 50319-0146, telephone number: 515-281-5295, or the Director of the Office for Civil Rights, U.S. Department of Education, Citigroup Center, 500 W. Madison Street, Suite 1475, Chicago, IL 60661-4544, telephone number: 312-730-1560, FAX number: 312-730-1576, TDD number: 877-521-2172, email: OCR.Chicago@ed.gov.

Contents

Executive Summary 4
Task Force Membership 5
Facilitator 5
Staff Support 5
Task Force Process 6
Iowa Science Standards 7
Implications for Assessment 8
Options for Assessments Aligned to the Iowa Science Standards 9
ACT Aspire Science Assessment 10
Match to the Task Force Rubric 10
Alignment to Iowa Science Standards 11
College or Career Readiness 12
Task Force Recommendations 14
Rationale 15
1. A Short-Term Solution 15
2. Grade Levels for Statewide Science Assessment 15
3. Statewide Assessment of Science 16
4. Grades Assessed 17
5. Funding for the New Assessments 17
6. Review of Science Assessment Options 17
7. Pursuit of Science Assessment Options Aligned with the Iowa Science Standards 17
Dissent 19
Appendices 20
Appendix 1: Current Statewide Science Assessment in Iowa 21
Appendix 2: Rubric and Additional Survey Questions 22
Appendix 3: Sample Performance Expectation 47
Appendix 4: Depth-of-Knowledge Explanation and Examples 48
Appendix 5: Resources 55
Appendix 6: Iowa Code Section 256.7(21)(b) 56
Appendix 7: Task Force Meeting Schedule 57

Executive Summary

In August 2015, the Iowa State Board of Education directed the Iowa Assessment Task Force – which was legislatively charged with recommending a statewide assessment of student performance for accountability purposes – to reconvene to recommend a statewide science assessment. Comprised of practicing teachers and administrators, technical assistance and professional development providers, higher education, and one representative each from the Iowa Department of Education, the Iowa Business Council, and a parent, the Task Force met over a seven month period to study the issues and opportunities around science assessment and to deliberate what is best for Iowa's children.

After careful study and deliberation, the Task Force makes the following recommendations for the statewide science assessment.

The Task Force Recommends:

1. **A Short-Term Solution**
   Through a unanimous vote in December 2015, the Task Force adopted the motion that the Task Force will recommend a short-term assessment that will be used no longer than the 2019-20 school year.
2. **Grade Levels for Statewide Science Assessment**
   Through a unanimous vote in December 2015, the Task Force adopted the motion that the Task Force will recommend that the science assessment be administered once in each of the three grade spans (3-5, 6-8, 9-12) in the short term.
3. **Statewide Assessment of Science**
   Through a vote of 13 to 1, the Task Force recommends the ACT Aspire science assessment be used as the statewide science assessment that will be used no longer than the 2019-20 school year, to be administered once in each of the three grade spans 3-5, 6-8, and 9-12.
4. **Grades Assessed**
   Through a unanimous vote, the Task Force recommends that the grades assessed in science be 5, 8, and 10.
5. **Funding for the New Assessments**
   Through a unanimous vote, the Task Force recommends the state appropriate funds to provide all districts access to the ACT Aspire science assessments and supports.
6. **Review of Science Assessment Options**
   Through a unanimous vote, the Task Force recommends the Task Force should meet at least annually through 2020 to review science assessment options.
7. **Pursuit of Science Assessment Options Aligned with the Iowa Science Standards**
   Through a vote of 13 to 1, the Task Force recommends that the Iowa Department of Education pursue additional options for accessing a statewide accountability assessment aligned to the Iowa Science Standards, such as developing an Iowa-only assessment or asking other states if they would join a consortium to share costs and expertise to develop an assessment.

Task Force Membership

Ann Benzshawel[1], Secondary Curriculum Facilitator, Des Moines Public Schools, Des Moines
Catherine Blando, Supervisor for Student Teachers, Iowa Wesleyan University, Cedar Rapids
Kathy Brenny, Consultant, Prairie Lakes Area Education Agency, Storm Lake
Martha Bruckner, Superintendent, Council Bluffs Community School District, Council Bluffs
Joe DeHart, Executive Director, Institutional Effectiveness, DMACC, Ankeny
Lowell Ernst, Director of K-12 Instruction, Pella Community School District, Pella
Harry Heiligenthal, Leadership Development Director, Iowa Association of School Boards, Des Moines
Tina Hoffman, Regional Administrator, Grant Wood Area Education Agency, Cedar Rapids
Mary Jackson[2], Resource Manager, Iowa Vocational Rehabilitation Services, Des Moines
Mark Lane, Director of Human Resources, Urbandale Community School District, Urbandale
Jo Ellen Latham, Director of Curriculum and Instruction, Southeast Polk Community School District, Pleasant Hill
Jon McKenzie, Director of Assessment & Comprehensive Improvement, Area Education Agency 267, Cedar Falls
Angela Olson, Associate Principal, Xavier High School, Cedar Rapids
Elliott Smith, Executive Director, Iowa Business Council, Des Moines
David Tilly, Deputy Director, Iowa Department of Education, Des Moines
Denise Wall, Educator, IKM-Manning Middle School, Manning
Tammy Wawro, President, Iowa State Education Association, Des Moines
Melanie Wirtz, Teacher, Peet Junior High School, Cedar Falls
Karen Woltman, Parent, North Liberty

Facilitator

Circe Stumbo, President, West Wind Education Policy Inc., Iowa City

Staff Support

Colleen Anderson, Iowa Department of Education, Des Moines
Diane Chadwick, Iowa Department of Education, Des Moines
Tom Deeter, Iowa Department of Education, Des Moines
Connor Hood, Iowa Department of Education, Des Moines
Kris Kilibarda, Iowa Department of Education, Des Moines

[1] Replaced member Shelly Bosovich.
[2] Replaced member Ruth Allison.

Task Force Process

Immediately after the new Iowa Science Standards were adopted, the Task Force reconvened, using the same norms and agreements and decision-making rules the Task Force established in November 2013. During September and October, the Task Force studied and discussed the newly-adopted Iowa Science Standards, particularly deliberating implications for assessment.

On November 3, 2015, the Task Force issued an initial Request for Information (RFI) on science assessment to gather information about current and future availability of science assessments designed to measure the Iowa Science Standards. The following four vendors responded to the initial RFI:

- ACT Aspire;
- edifyAssess, submitted by Iowa School Finance Information Services;
- Measured Progress; and
- WestEd.

The Task Force also requested information about science assessments or assessment plans of other states that have adopted the Next Generation Science Standards (NGSS). The Center on Standards and Assessment Implementation collected this information.

After reviewing assessment possibilities at the November meeting, the Task Force decided to update the rubric and additional survey questions that were developed for the Task Force's review of mathematics and reading assessments. The updated rubric and additional survey questions were approved at the December 15, 2015, meeting (see Appendix 2).

Also at the December meeting, the Task Force reviewed and discussed excerpts from *Developing Assessments for the Next Generation Science Standards* published by the National Academies Press. Following deliberation, the Task Force adopted two motions: that the Task Force would recommend a science assessment for the short term and that the Task Force would recommend administration of the science assessments at one grade level in each of three grade spans in the short term.

A second RFI was issued December 18, 2015, inviting submission of the information requested in the updated rubric and additional survey questions. Only one vendor submitted a response: ACT Aspire.

The ACT Aspire submission was scored by two small groups. The ACT Aspire submission and small group scores were discussed at the January meeting and the Task Force determined that ACT Aspire should be invited for a vendor interview.

ACT Aspire was interviewed at the February meeting and provided additional materials, including operational forms and sample score reports, for Task Force review at that time.

The Task Force deliberated recommendations after the interview and again at the March meeting. The final vote on the recommendations in this report was held at the March 11, 2016, meeting.

Iowa Science Standards

Iowa has had statewide science standards since 2008. Pursuant to Executive Order 83, the statewide science standards were reviewed by the Science Standards Review Team in 2014-15. Upon completion of their review, the Science Standards Review Team recommended that the Next Generation Science Standards (NGSS) performance expectations be adopted as Iowa's new science standards, with modifications described below. In August 2015, the State Board voted unanimously to adopt the Science Standards Review Team's recommendation.

The NGSS, finalized in April 2013, were developed by twenty-six lead partner states, including Iowa, in collaboration with the national nonprofit organization Achieve Inc.[3] Development of the standards was guided by the National Research Council's conceptual framework, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*.[4] Sixteen states and the District of Columbia have adopted the NGSS to date.[5]

The NGSS are organized as performance expectations that describe what students should be able to do to demonstrate mastery of the science standards at each grade level. The performance expectations offer a focused and coherent learning progression in science over the course of a student's K-12 experience and are aligned with the Common Core Mathematics and English Language Arts Standards.

The performance expectations incorporate the three, equally important NGSS dimensions: Disciplinary Core Ideas, Crosscutting Concepts, and Science and Engineering Practices.

Disciplinary Core Ideas are the key content knowledge in the disciplines of earth and space science, life science, physical science, and technology/engineering. **Crosscutting Concepts** are unifying themes across disciplines, such as patterns or cause and effect, that can help students build understanding of disciplinary core ideas. **Science and Engineering Practices** are the knowledge-specific skills used in engaging in scientific investigations or engineering solutions to identified problems.

Performance expectations are not instructional or assessment tasks and do not represent the only way that Disciplinary Core Ideas, Crosscutting Concepts, and Science and Engineering Practices can be integrated for instructional and classroom assessment purposes.

States can adopt either the NGSS performance expectations or the performance expectations plus the foundational boxes outlining the material from each of the three dimensions incorporated into each of the performance expectations as their state standards (see Appendix 3). Upon the recommendation of the Science Standards Review Team, the State Board adopted only the performance expectations as the Iowa Science Standards in order to emphasize the need for integrated, three-dimensional instruction that aligns with the standards.

---

[3] See http://www.nextgenscience.org/.

[4] National Research Council. (2012). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. http://nap.edu/13165

[5] The other fifteen states are Arkansas, California, Connecticut, Delaware, Hawaii, Illinois, Kansas, Kentucky, Maryland, Nevada, New Jersey, Oregon, Rhode Island, Vermont, and Washington.

Additionally, the State Board adopted the performance expectations as grade-level specific, rather than as grade span standards, for the middle school grades (6-8). This is intended to achieve consistency across schools and districts, ensure integration of each of the four science disciplines at each grade level K-8, and facilitate collaboration.

The Iowa Science Standards are to be fully implemented for the elementary and high school grades and one grade at the middle school level by the end of the 2018-19 school year. Districts will have two additional years to fully implement the middle school grade standards, with full implementation in all grade levels required by the 2020-21 school year.

Implications for Assessment

Adoption of the NGSS performance expectations as the Iowa Science Standards have several important implications for statewide science assessment. New assessments aligned to the current Iowa Science Standards will be needed.

What would (or could) these new statewide assessments, specifically designed to measure three-dimensional science standards, look like? The Task Force spent substantial time studying and discussing this question. The short answer: probably very different from traditionally available, large-scale science assessment.

The Iowa Science Standards are written as performance expectations, not separate lists of science content to be covered and skills to be acquired, to emphasize that the doing of science (science and engineering practices) is as important as knowing science content (disciplinary core ideas and crosscutting concepts). Furthermore, not only are the three dimensions of equal importance, but they must be integrated. It is intended that students will develop deep understanding of the disciplinary core ideas through engaging in multiple science and engineering practices and using various crosscutting concepts as lenses through which to consider scientific phenomena; in short, students will learn science through doing science.

This is a critical conceptual shift, the importance of which cannot be overstated: In order to be fully aligned, new assessments will need to be able to measure what students are able to do and what they know in science in an integrated fashion, which means selected response test items alone will not be sufficient; performance tasks will be required to fully measure the Iowa Science Standards.

The Task Force spent some time discussing ideas about possible features of NGSS-aligned science assessment in the long term, but ultimately concluded that whatever the future holds, currently available assessments are not there yet and the assessments chosen for the short and middle term should put us on the path to the statewide science assessments Iowa will need in the long term.

One other important implication for assessment is that there may be opportunities for Iowa to collaborate in the development of new science assessments because other states have also adopted the NGSS.

Options for Assessments Aligned to the Iowa Science Standards

As part of the review process, the Task Force collected information about assessments designed to measure student achievement toward the Iowa Science Standards/NGSS. Information was collected through the RFI process in addition to a scan of what NGSS states are doing for their science assessment conducted by the Center on Standards and Assessment Implementation. Four vendors responded to the RFIs. Iowa Testing Programs, the entity that created and administers the current Iowa Assessments, did not submit a response.

Of the four responses to the initial RFI, only one, from ACT Aspire LLC, involved a currently operational assessment. One vendor, Measured Progress, expressed interest in partnering with the Iowa Department of Education but offered no specific assessment option for review.

Two vendors described work creating test item banks and proposed to build statewide science assessments for Iowa based on their existing item banks. The edifyAssess response proposed a pilot test in 2016-17, with an operational assessment available for administration in the spring of 2018. The WestEd response, based on their work as vendor for the Council of Chief State School Officers' Science Assessment Item Collaborative,[6] proposed working with Iowa, either individually or as part of the Collaborative, to develop aligned assessments that could be operational as soon as the spring of 2018.

In addition to requesting information from assessment vendors, the Task Force asked the Midwest Comprehensive Center if they could determine what other states were doing to have aligned assessments. They worked with the national Center on Standards and Assessment Implementation, who scanned state department of education websites to determine how states would be measuring science achievement. The Center on Standards and Assessment Implementation found that a few states are pilot testing NGSS-aligned assessments this school year, while most of the NGSS states are administering existing science assessments; some of these states have plans for developing NGSS-aligned science assessments. One state, Arkansas, has adopted the ACT Aspire.

In short, while some vendors and states are planning to develop or are currently in the process of developing assessments aligned to the NGSS, no currently operational assessments were developed specifically to measure the Iowa Science Standards/NGSS.

[6] Fourteen states are members of the Council of Chief State School Officers' Science Assessment Item Collaborative: Arkansas, California, Connecticut, Delaware, Illinois, Kentucky, Hawaii, Maryland, Massachusetts, Michigan, Nevada, Oregon, Washington, and West Virginia.

ACT Aspire Science Assessment

The only currently operational assessment offered for Task Force review is the ACT Aspire science assessment. ACT Aspire is a joint venture of ACT Inc. and Pearson Inc. and offers assessments for grades 3-10 in five subject areas: English, math, reading, science, and writing. The ACT Aspire assessments were developed to be modular; each subject area assessment is designed to be capable of being administered as a standalone assessment. The ACT Aspire assessments are available for use in the 2016-17 school year.

ACT Aspire science assessments are fixed-form tests that can be administered on computers or in a paper-and-pencil format. The ACT Aspire science assessments include selected response, technology-enhanced (computer-based administration only), and constructed response item types. Selected response items require a student to choose an answer from options provided by the test writers. Technology-enhanced items make use of technology to add either to the item prompt (video or audio stimuli instead of, or in addition to, reading passages, for example) or in recording answers (having students drag and drop items on screen or using drawing tools to create a graph, instead of marking a multiple-choice option). Constructed response items require students to write or create their own answers (for example, writing a sentence or paragraph). Through a combination of these test item types, ACT Aspire science assessments cover Norman Webb's Depth of Knowledge[7] levels 1-3.

ACT Aspire science assessments are timed tests and take fifty-five minutes to administer to students in each of the grade levels three through ten. ACT Aspire provided cost estimates for administration of the science assessment of $8.00 per student for computer-based administration and $12.00 per student for administration in a paper-and-pencil format. There would be an additional, estimated cost of $1.20 per student for five-page student score reports printed in color.

Match to the Task Force Rubric

Iowa Code Section 256.7(21)(b)(2) outlines the minimum legislative requirements for statewide assessments administered in the 2016-17 school year and beyond. These legislative requirements formed the foundations of the Task Force rubric and additional survey questions (see Appendix 2).

Through the scoring process, it was determined that ACT Aspire met or exceeded the requirements for fairness, availability, piloted and tested in Iowa, career and college readiness, and the sub-criterion related to the availability of training on assessments and interpretation of results (technical supports). Based on the Task Force review of ACT Aspire's materials prior to the interview, their assessments scored less well on validity and reliability, but were deemed technically adequate; additional evidence of validity and reliability was provided at the interview.

ACT Aspire did not meet the requirements for accurately describing achievement and for alignment, which was expected because both of these criteria relate to measuring and reporting on the Iowa Science Standards and the ACT Aspire was developed prior to the development and adoption of the Iowa Science Standards. This will be discussed further in the following subsection, however, it

[7] Depth of Knowledge refers to cognitive complexity of a task. Norman Webb's formulation allows for tasks to be described at four levels of cognitive complexity, starting with Level 1 – "recall and reproduction" – and going through Level 4 – "extended" thinking." For a more complete description of DOK, please see Appendix 4.

should be noted here that there were no assessments available for Task Force review that would have scored well on these criteria.

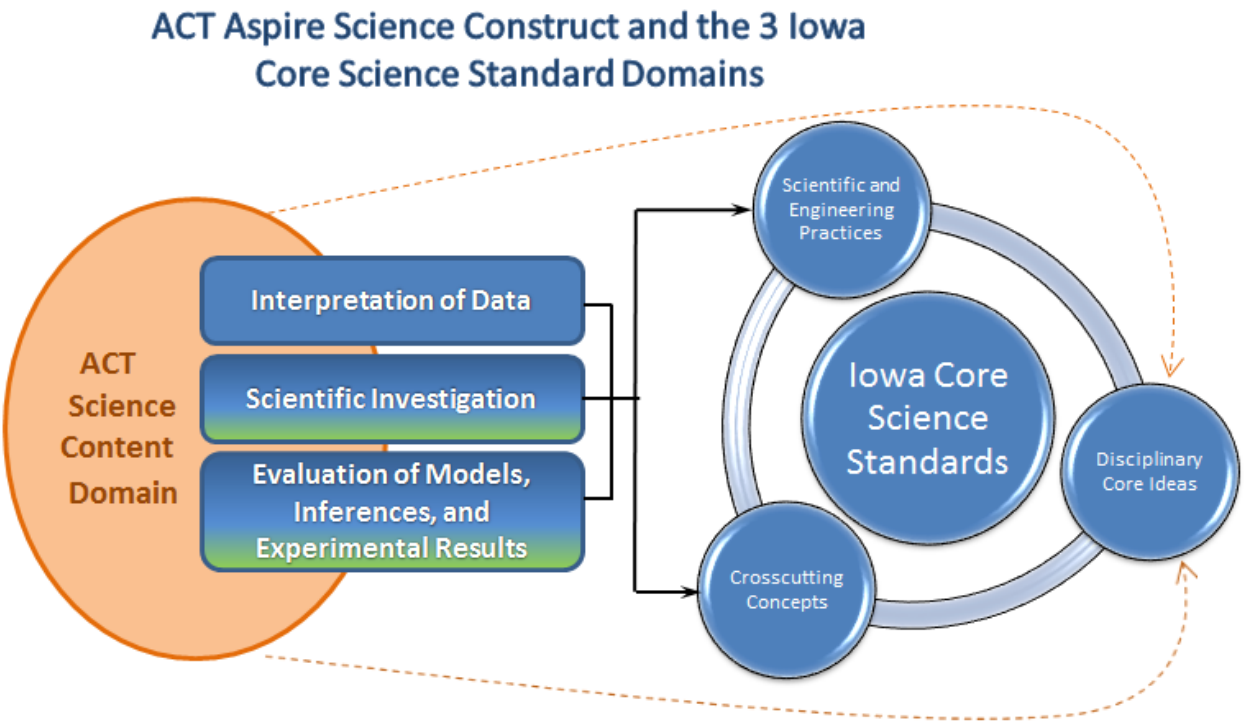
Alignment to Iowa Science Standards

The key feature of any assessment is its alignment to the standards, meaning the extent to which a test measures what we want students to know and be able do at each grade level assessed. Ideally, alignment to a desired set of standards would be part of the assessment development process, though sometimes post hoc alignment studies are done to show the level of alignment between an already developed assessment and other standards.

Our expectations for what students should know and be able to do at each grade level are contained in the Iowa Science Standards. A statewide science assessment would, ideally, be in complete alignment with these standards.

The ACT Aspire science assessments were developed to provide measurement and predictions based on the ACT College and Career Readiness Standards; the ACT Aspire science assessments were not developed specifically to measure the Iowa Science Standards/NGSS.

In response to the initial RFI, ACT Aspire provided an internal comparison study demonstrating areas of alignment between ACT Aspire science assessments and the Iowa Science Standards/NGSS. In response to the second RFI, ACT Aspire described the science assessments as “focus[ing] on science process skills (Science and Engineering Practices and Cross-Cutting Concepts) assessed in content-rich, authentic scientific scenarios (Disciplinary Core Ideas).” ACT Aspire also provided the following graphical representation of the relationships between the three ACT Aspire science subscore reporting domains and the three dimensions of the Iowa Science Standards.



While the post-hoc alignment analysis provided by ACT Aspire demonstrates that educators can find ways to discern alignment between the Iowa Science Standards' Disciplinary Core Ideas and the ACT Aspire assessment, ACT Aspire does not report subscores in any of the four disciplines in the Disciplinary Core Ideas. ACT Aspire indicated reporting subscores in each of the four disciplines would not be feasible as "[i]t is difficult to estimate coverage of disciplinary core ideas as ACT Aspire represents a large variety of content sequences within a domain sampling model."

The science subscore reporting categories are also imperfectly aligned with regard to the Iowa Science Standards/NGSS science practices. ACT Aspire claims that all eight of the Iowa Science Standards/NGSS science practices are represented within the three science domains used by ACT Aspire to report science subscores. During the interview, ACT Aspire further explained that, due to the length of the science assessments, there would not be enough questions in each of the eight science practices to reliably report performance in eight separate categories. Therefore, they chose to report subscores in only three separate domains on three of the eight practices.

Another area of imperfect or non-alignment addressed by ACT Aspire is the lack of performance tasks in their assessment. At the interview, ACT Aspire noted that assessing DOK 4 would require students to engage in performance tasks over several days if not weeks, which would not be feasible given their assessment design. ACT Aspire contends that even without performance task items, the science assessment is designed to yield valuable information to students and educators.

A final area of imperfect or non-alignment, also noted by ACT Aspire, is engineering practices and terminology, which are not directly assessed on the ACT Aspire science assessments. ACT Aspire explained at the interview that engineering practices are not directly assessed on the ACT Aspire science assessments because they have not yet been identified as important to success in first year college courses or careers. ACT Aspire further noted that assessment of engineering practices would be added to ACT Aspire assessments in the future if national curriculum survey data were to indicate that engineering practices have become important to college and career readiness.

College or Career Readiness

Another desired feature of statewide assessment is a college or career readiness indicator. College or career readiness addresses the ability of a test to predict future success after completion of high school. Definitions of these concepts are still evolving, but the most common current definition of college readiness involves predicting later success.

ACT has developed its own college and career readiness standards, as previously mentioned. These standards were developed and are periodically updated by ACT based on information collected by ACT in national curriculum surveys which ACT conducts every three years. Using survey responses, ACT identifies the "non-negotiable skills" needed for success in first year college courses or careers. The "non-negotiable skills" are then used to determine content and skills to be tested on the ACT. The ACT Aspire science assessments are based on the ACT College and Career Readiness Standards and are developed by working backwards from the "non-negotiable skills" identified by ACT.

The ACT Aspire science assessments report an overall science score in addition to subscores in three science domains: Interpretation of Data; Scientific Investigations; and Evaluating Models, Inferences, and Experimental Results.

In addition to ACT readiness levels or benchmarks at each grade level, the ACT Aspire science assessments reports also include a predicted ACT science score for students in grades nine and ten.

Task Force Recommendations

Ideally statewide science assessments would be in perfect alignment with the Iowa Science Standards. No currently operational science assessments meet this ideal, though, based on our review, we expect specifically aligned science assessments to be developed within the next few years. In the meantime, however, state and federal law requires administration of annual statewide science assessments to Iowa students. Therefore, the Task Force makes the following recommendations for statewide assessment of student progress in science.

1. **A Short-Term Solution**

   Through a unanimous vote in December 2015, the Task Force adopted the motion that the Task Force will recommend a short-term assessment that will be used no longer than the 2019-20 school year.

2. **Grade Levels for Statewide Science Assessment**

   Through a unanimous vote in December 2015, the Task Force adopted the motion that the Task Force will recommend that the science assessment be administered once in each of the three grade spans (3-5, 6-8, 9-12) in the short term.

3. **Statewide Assessment of Science**

   Through a vote of 13 to 1, the Task Force recommends the ACT Aspire science assessment be used as the statewide science assessment that will be used no longer than the 2019-20 school year, to be administered once in each of the three grade spans 3-5, 6-8, and 9-12.

4. **Grades Assessed**

   Through a unanimous vote, the Task Force recommends that the grades assessed in science be 5, 8, and 10.

5. **Funding for the New Assessments**

   Through a unanimous vote, the Task Force recommends the state appropriate funds to provide all districts access to the ACT Aspire science assessments and supports.

6. **Review of Science Assessment Options**

   Through a unanimous vote, the Task Force recommends the Task Force should meet at least annually through 2020 to review science assessment options.

7. **Pursuit of Science Assessment Options Aligned with the Iowa Science Standards**

   Through a vote of 13 to 1, the Task Force recommends that the Iowa Department of Education pursue additional options for accessing a statewide accountability assessment aligned to the Iowa Science Standards, such as developing an Iowa-only assessment or asking other states if they would join a consortium to share costs and expertise to develop an assessment.

Rationale

1. A Short-Term Solution

Task Force will recommend a short-term assessment that will be used no longer than the 2019-20 school year.

The Iowa Science Standards represent a major conceptual shift in K-12 science education. Focused on having students learn through "doing what scientists and engineers do" (that is, evaluate and extend thinking and solve problems), the Iowa Science Standards clearly cannot be fully assessed without extended constructed response and performance task items. As teachers and students transition to the new framework and vocabulary of the Iowa Science Standards, our statewide assessments and student performance score reports must also transition to using the same frameworks and vocabulary.

Our scan of currently available assessments demonstrated that large-scale assessments in science with extended constructed response and performance task items and using the language of the Iowa Science Standards/NGSS are not yet available. While new, aligned assessments are being developed, it is critical that science continues to be assessed at the state level. Iowa students, families, educators, and policy-makers need information about the effectiveness of science instruction and student needs. Further, the federal government requires that science achievement be reported on an annual basis; Iowa would risk losing federal education funding if science assessments were not administered in the interim.

Because we understood that no assessment specifically designed to measure the Iowa Science Standards/NGSS would be available for review, we determined that any assessment recommendation should be made for the short term only and not to be seen as adequate to meet Iowa's long term science assessment needs. Until new assessments are available, Iowa should administer science assessments that are as aligned as much possible to the Iowa Science Standards, however imperfectly, and that also provide useful information to Iowa's students, families, educators, and the broader community.

Aligned assessments are expected to be available no later than the 2019-20 school year.

2. Grade Levels for Statewide Science Assessment

The Task Force will recommend that the science assessment be administered once in each of the three grade spans (3-5, 6-8, 9-12) in the short term.

Until the field has developed assessments that provide multicomponent tasks that assess all three dimensions in the science standards, the National Research Council's Committee on Developing Assessments of Science Proficiency in K-12 recommends that large-scale assessments for accountability should be administered at least once, but no more than twice, in each grade span.

This matches federal requirements; according to the federal Elementary and Secondary Education Act, science assessments are required at least once in each of three grade spans, 3-5, 6-8, and 9-12. We believe it also matches best practice at a time when better-aligned assessments are not

available. While science is critically important, it is not worth the kind of student time it would take to assess science in every single grade if the assessment is currently so imperfect.

Iowa also took the somewhat unique position that science in the middle grades should be taught in a specific sequence that is different from what is called for in the NGSS. Therefore, it is likely that any assessment not specifically designed for Iowa or flexible enough to allow for content to be assessed at all three middle-level grades would not work for Iowa, unless it were only administered at the 8th grade level. By then, all of Iowa's students will have been exposed to the same content as students in other states. Therefore, only assessing student achievement in grades 6-8 one time, in grade 8, would work well for Iowa.

However, Iowa Code currently requires assessments in science at every grade 3-11. Therefore, in December 2015, prior to sending out our second Request for Information from assessment vendors, we voted to recommend that while using a statewide assessment in the short-term, the state should only assess one grade in each of the grade spans (3-5, 6-8, and 9-12). We built this into our scoring rubric, offering a score of "above good" for any assessments that had an 11th grade option as part of their package.

3. Statewide Assessment of Science

The Task Force recommends ACT Aspire science assessment be used as the statewide science assessment that will be used no longer than the 2019-20 school year, to be administered once in each of the grade spans 3-5, 6-8, and 9-12.

We are concerned about the areas in which the ACT Aspire are only partially, or not at all, aligned to the Iowa Science Standards, as well as with the difficulty we had in interpreting the subscore reporting domains in terms of the Iowa Science Standards. We also considered the extent to which the assessment accurately describes student achievement and growth. The key in this criterion is that the three dimensions (disciplinary core ideas, scientific and engineering practices, and crosscutting concepts) need to be present, and the assessment is built around integrated ways to measure them together. Further, though ACT Aspire demonstrated strong vertical alignment between the ACT Aspire assessments, the ACT, and college and workplace training readiness, they were not able to demonstrate vertical alignment with the learning progressions built into the Iowa Science Standards. This is important because these learning progressions are well articulated in the Iowa Science Standards and can help understand a student's growth in science understanding.

However, all of this was expected, as the assessment was developed prior to adoption of the Iowa Science Standards. While ACT Aspire is not perfectly aligned with the Iowa Science Standards, we knew that no existing assessments would be in perfect alignment. As we wait for new assessments to be developed, ACT Aspire reports offer useful information to students, families, and educators, including a predicted ACT science score for students in grades nine and ten.

Further, ACT Aspire can be administered, and is used by other districts already, in grade spans, which matches our previous recommendation. This is significant to our recommendation to adopt ACT Aspire as the short-term science assessment. ACT Aspire is available for grades 9 and 10, but if we need to have an assessment at grade 11, it would have to be the ACT. It is not clear that ACT science could be taken as a standalone assessment and the current price for the complete ACT (without writing) is $39.50 per student. This cost would be prohibitive, making the choice to assess only once in grade 9 or 10 even more attractive.

One member dissented from this Task Force recommendation; the dissent is provided at the conclusion of all the recommendations.

4. Grades Assessed

The Task Force recommends that the grades assessed in science be 5, 8, and 10.

Educators, students, and families need to know the grade levels to be assessed as soon as possible. This recommendation is to assess at the highest possible grade within the grade spans to give students the most opportunity to learn the material as possible, while understanding that high school students will not have four years to access the material, but there is no ACT Aspire assessment available for grades 11 or 12.

5. Funding for the New Assessments

The Task Force recommends the state appropriate funds to provide all districts access to the ACT Aspire science assessments and supports.

It is critical to maximize the state's investment to use the ACT Aspire resources to improve teaching and learning, in addition to meeting the state's accountability needs. The state should provide funding for the assessments themselves and for scoring, as well as for interim assessments and stakeholder communications. The state should work with teacher leaders to help support K–12 teachers in their use of the formative assessment processes to adjust teaching to improve student learning.

6. Review of Science Assessment Options

The Task Force recommends the Task Force should meet at least annually through 2020 to review science assessment options.

While no currently operational assessments were designed specifically to measure the Iowa Science Standards/NGSS, some vendors and states are planning to develop or are currently in the process of developing assessments aligned to the NGSS. Because alignment is a critical feature of an assessment, Iowa Science Standards/NGSS-aligned assessments should be reviewed annually to ensure that a high-quality Iowa Science Standards/NGSS aligned assessment can be implemented in Iowa as soon as possible.

7. Pursuit of Science Assessment Options Aligned with the Iowa Science Standards

The Task Force recommends that the Iowa Department of Education pursue additional options for accessing a statewide accountability assessment aligned to the Iowa Science Standards, such as developing an Iowa-only assessment or asking other states if they would join a consortium to share costs and expertise to develop an assessment.

Because an Iowa Science Standards-aligned assessment is of critical importance to Iowa's students and teachers, and because there are no federal funds currently available to fund development of aligned assessments, the Iowa Department of Education should actively pursue science

assessment development options to ensure high-quality, aligned science assessments are available for implementation in Iowa as soon as possible. Such science assessments could consist purely of statewide, on-demand exams, or they potentially also could incorporate classroom-embedded assessments, which would allow for student engagement in all three dimensions of science over time. They could be administered in every grade 3-11 or they could be administered once in each of several grade spans. They could assess every child or they could use a matrix sampling approach to evaluating system performance. All of these questions should be considered as the Iowa Science Standards are being implemented.

Rather than waiting for NGSS-aligned assessments to be developed by external developers, Iowa should be actively involved in the development process to ensure the ultimate assessments align with our needs and expectations. We could achieve this goal by developing our own unique assessment, but such an approach would be costly. With sixteen other states adopting the NGSS, the possibility for a collaboratively-developed assessment in which multiple states share the cost of development is very real. It would behoove Iowa to look into forming a consortium with other states to develop NGSS-aligned assessments.

One member dissented from the seventh Task Force recommendation, but did not write a dissent.

Dissent

One member dissented from the third Task Force recommendation:

The intent and purpose of the ACT Aspire Science Assessment is to predict whether or not students will be successful in college and career level science. The design and purpose of Iowa's state assessment in science should be to measure achievement and growth based on the Next Generation Science Standards (NGSS). ACT's Technical Bulletin acknowledges "while considerable alignment between ACT Aspire and the NGSS does exist, the ACT Aspire assessments are not based on the NGSS, and they do not directly assess the NGSS" (ACT, 2014). ACT Aspire assessments are only available in grades 3-10. Particularly at the high school level, students will be tested on science content they have not yet been taught. Therefore, 10th grade students taking ACT Aspire will be held accountable for content they may be on a pathway to learn in their 11th and 12th grade years of school. When the assessment is used to measure achievement rather than to predict later success, the results may be misinterpreted as poor performance by the teacher, the student, and the school district, when in fact, it is the result of an assessment that is not aligned to the science standards.

The Next Generation Science Standards were adopted into the Iowa Core on August 6, 2015. Performance Expectations in the Iowa Core indicate what students should be able to do in order to demonstrate they have met the standard, thus providing clear and specific targets for curriculum, instruction, and assessment. In the ACT Aspire Technical Bulletin #1, they state that "some areas of achievement are not measured on ACT Aspire summative assessments. Standards that require extended time, advanced uses of technology, active research, collaboration, producing evidence of a practice over time, or speaking and listening are not currently assessed" (ACT, 2014).

Engineering is a strong component of the Iowa Core Science standards, and ACT Aspire does not measure this separately but instead includes it with science practices. According to the NGSS overview document, "although engineering design is similar to scientific inquiry, there are significant differences. For example, scientific inquiry involves the formulation of a question that can be answered through investigation, while engineering design involves the formulation of a problem that can be solved through design" (NGSS, 2013). It will be important for the state of Iowa to include engineering practices in the next science assessment in order to prepare students for a future that embraces engineering aspects in both career opportunities and everyday life.

While I realize the available science assessments are limited, and that this is a temporary recommendation, I cannot support holding students, teachers, and schools accountable with an assessment that does not measure what we are asking them to teach and learn. I respectfully dissent from the State of Iowa's Assessment Task Force's recommendation of ACT Aspire as the state science assessment.

ACT, 2015. *ACT Aspire Technical Bulletin #1: Summative Assessment*. http://www.discoveractaspire.org/pdf/2014_ACT-AspireTechnicalBulletin1.pdf)

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Appendices

Appendix 1: Current Statewide Science Assessment in Iowa
Appendix 2: Rubric and Additional Survey Questions
Appendix 3: Sample Performance Expectation
Appendix 4: Depth of Knowledge
Appendix 5: Resources
Appendix 6: Iowa Code Section 256.7(21)(b)
Appendix 7: Task Force Meeting Schedule

Appendix 1: Current Statewide Science Assessment in Iowa

The current statewide science assessments are the Iowa Assessments, formerly known as the Iowa Tests of Basic Skills and the Iowa Tests of Educational Development, which were developed and are administered by Iowa Testing Programs. The Iowa Assessments in science are fixed-form tests, administered largely in paper-and-pencil format, with an online administration option offered since the 2014-15 school year. The entire test is made up of selected response (multiple-choice) items.

The Iowa Assessments are timed tests, with the science test portion taking thirty-five minutes to administer to students in each of the grades three through eight, and forty minutes to administer to students in each of the grades nine through twelve.

The current cost of the Iowa Assessments is $4.25 to $6.25 per student for paper-and-pencil tests or $13.00 per student for online tests for basic scoring and reporting services.

The current Iowa Assessments provide a statistical prediction of college readiness based on the ACT college readiness benchmark scores.

Appendix 2: Rubric and Additional Survey Questions

Iowa Assessment Task Force — Science Assessment Review Rubric

1. Fairness

Please demonstrate how you assure fairness in your assessments.

1. Statistical evidence of fairness in development
2. Fairness in development (Universal Design)
3. Fairness in administration; accommodations (accessibility)
4. Fairness in administration; standardized directions
5. Fairness in administration; practice items

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Statistical evidence of fairness in development	Differential Item Functioning Analysis (DIF) used to refine items on test, with appropriate sample	Test development included all subgroups (with appropriate sample and stratified random sample)	Test development population was diverse but not specified further	Little or no evidence of empirical analysis of fairness/bias provided	How did the developer attend to bias using statistical methodology? Did they at least include all subgroups while testing the function of items? Did they directly examine item function for each subgroup (DIF analysis)?
Fairness in development (Universal Design)	Process described for reviewing item wording, visuals for accessibility, cultural bias, offensive content, visual distractors in items and directions (all present) Exemplars provided	Process described for reviewing item wording, visuals for accessibility, cultural bias, offensive content, visual distractors in items and directions (majority present) Exemplars provided	Items and directions reviewed by experts for bias/fairness, specific details not provided	No evidence of expert review of fairness/bias provided	What does the developer tell us about processes used to examine and minimize bias? There are a variety of standard methods used while developing and refining items, tests and other materials. These are often known as principles of universal design. These help insure accessibility of test items.

Fairness in administration; accommodations (accessibility)	Validated, feasible accommodations are described, including for example, sensory impairments, limited English proficiency	Feasible and appropriate accommodations are described, including for example, sensory impairments, limited English proficiency	List of accommodations	No evidence of guidance for accommodations	Does the test include specific information about acceptable accommodations for students who may have disabilities such as visual or hearing impairments where feasible and appropriate? Validation means that there have been studies of the effects of these accommodations, with appropriate cautions about risks to data use.
Fairness in administration; standardized directions		Administration is standardized		No evidence of standardized administration	This is a yes/no item. Either administration directions are standardized, or they are not. Standardized instructions help insure fairness in administration.
Fairness in administration; practice items	Practice items include sample questions that represent all item structures that will be present on the assessment at each grade. Samples of relevant accommodations presented when appropriate	Availability of practice items at each specific grade. Samples of relevant accommodations presented when appropriate	Practice test items for grade spans	No practice test content available	This is about the question structures - do the items create a barrier to measuring knowledge and skill? Practice items provide students with an opportunity to experience the item structures that will be present on the test. Accommodations are defined in practice items where appropriate.

2. Availability

Please demonstrate the availability of your assessments.

1. Grade availability
2. Availability in last quarter of school year

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Grade availability	Available grades 3-11	Available at grades 3-5, 6-8, and 9-12		Not available at all three grade bands	The Task Force recommends that the science assessment be administered once in each of the three grade bands (grades 3-5, 6-8, 9-12) in the short-term.
Availability in last quarter of school year		Yes		No	Legislation requires spring administration

3. Describes Achievement

Please demonstrate how your assessments accurately describe student achievement.

1. Accurately describes student achievement
2. Accurately describes growth

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Accurately describes student achievement	Materials indicate reporting addresses all three dimensions in an integrated fashion	Materials indicate all three dimensions are addressed with evidence of level of skill in addition to a percentage of mastery	Reports describe content relative to disciplinary core ideas and scientific and engineering practices, but they are isolated/not integrated.	Reports describe content. The three dimensions (if present) are represented as content items and not identified.	The three dimensions are (1) disciplinary core ideas, (2) scientific and engineering practices, and (3) crosscutting concepts. The key in this criterion is that the three dimensions need to be present, and the assessment built around integrated ways to measure them together.
Accurately describes growth	Vertical scaling of results tied directly to learning progressions in the Iowa Core Science Standards and allow for measuring above and below grade level and grade spans for high and low performing children	Vertical scaling of results tied directly to learning progressions in the Iowa Core Science Standards	Achievement results reported on a vertical scale interpretable across grades/years	No vertical scaling	The new idea here is that the vertical scaling of the results is easily linked to the vertical articulation of learning progressions built into the Iowa Core Science Standards. This is important because these learning progressions are well articulated in the Iowa Core Science Standards and can help understand a student's growth in science understanding.

4. Validity

Please demonstrate how you have determined your assessment is valid

1. Criterion validity coefficient (correlational evidence)
2. Quality of Validity evidence

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Criterion validity coefficient (correlational evidence)	>0.8	>0.70	>0.60	<0.6 or none reported	Does this test measure the same thing as other tests?
Quality of Validity evidence	Multiple studies using criteria from level 2 with different populations and including all applicable statistics	Validity evidence using correct methodology with reasonable samples (size and representativeness). Comparison measures used are of reasonable technical quality and measure the desired content/constructs	Validity evidence using correct methodology, but marginal samples (size and/or representativeness). Comparison measures used measure the desired construct, but are not "mainstream" assessments and the samples used for the comparison measure's technical data are not representative (a sample of convenience).	Correct methodology, but poor or dated sample	Correlation measures that estimate validity should use assessments that are themselves reliable and valid, and measuring the desired construct - in other words, we are comparing reading tests with reading tests. Less desirable data may be found where the assessment used to validate is a less-than robust measure or use samples that are far from representative in either population, or timeframe. Further evidence of validity may be found in the alignment items.

5. Reliability

Please demonstrate how you have determined that your assessment is reliable

1. Internal consistency (alpha, split half, marginal)
2. Stability over time (test retest, alternate form)
3. Scorer consistency (inter-rater agreement in some form) (if applicable)
4. Quality of reliability evidence

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Internal consistency (alpha, split half, marginal)	>0.9	>0.8	>0.7	<0.7	
Stability over time (test retest, alternate form)	>0.9	>0.8	>0.7	<0.7	
Scorer consistency (inter-rater agreement in some form)	>0.9	>0.8	>0.7	<0.7	If applicable (i.e., scoring of the test requires human or machine scoring of student constructed response, where scorer error could be a factor). If not applicable, this item will not count for or against in the score.
Quality of reliability evidence	Multiple studies using different populations and including all applicable reliability statistics. Includes correct methodology as in #2.	Reliability evidence using correct methodology with reasonable samples (size, relevance to Iowa students and, representativeness).	Reliability evidence based on limited or non-representative populations	Reliability evidence on previous or related version	We are looking for sufficiently thorough reviews of reliability, with all applicable statistics. Not only are the reliability values important, but it is important for us to make some judgment about the methods used to gather the results, and how those values were calculated. High values resulting from low quality research data are suspect. We want to be certain the values reported are meaningful and interpretable in the Iowa context.

6. Piloted/Tested in Iowa

Please demonstrate that your assessment has been piloted in Iowa

1. Piloted in Iowa (item tryout)
2. Tested in Iowa (field tested)

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Piloted in Iowa (item tryout)		yes		no	Piloted means item tryouts and small sample test groups Pilot Test: A stand-alone administration of test items, tools or a system, to evaluate how particular items function prior to a field test and operational use. The pilot test generally occurs with a sample of students that matches the
Tested in Iowa (field tested)		yes		no	Field Test means larger scale testing of a fully developed test Field Test: An administration of the field test to evaluate how the test functions prior to operational use. This generally occurs after a pilot test, using a significantly larger, more representative sample of students than a

7. Alignment

Please demonstrate that your assessment is aligned in the following ways

1. Tables of specifications
2. Coverage of disciplinary core ideas
3. Coverage of scientific and engineering practices
4. Coverage of crosscutting concepts
5. Evidence of integration among the three dimensions of the Iowa Core Science Standards
6. Language is consistent with the Iowa Core Science Standards

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Tables of Specifications		Tables of specifications for the developed or new test are provided.		No table of specifications provided	The tables of specifications contain the blueprint for test construction, including scientific and engineering practices, crosscutting concepts and disciplinary core ideas.
Coverage of Disciplinary Core Ideas		All disciplinary core ideas are addressed at the discipline, core idea, and sub-idea level	Content coverage represented in the language of the Iowa Core Science Standards. Descriptions at the discipline and core idea level	Content is mapped to the Iowa Core Science Standards language. Descriptions at the discipline level.	This is the “granularity” of the content: Discipline = physical science, earth science, etc. Core Idea = e.g., matter and its interactions, etc. Sub-idea = e.g., structure and properties of matter, etc. *The core ideas are closest to what we are familiar with for science content.*

Coverage of Scientific and Engineering Practices		Practices represented in the language of the Iowa Core Science Standards. Descriptions reflect grade level progression	Practices in the Iowa Core Science Standards language are present	Practices not included or simply represented as inquiry	This is the application of the engineering and scientific practices. In the old world, if addressed at all, they were often addressed as if the “practice” was content or limited to a broad brushstroke of “inquiry.”
Coverage of Crosscutting Concepts			Crosscutting concepts in the Iowa Core Science Standards language are present	Crosscutting concepts are not included	This is probably new territory for most assessments. Although elements may have been present, attending to them directly is relatively new to the Iowa Core Science Standards.

Integration	The assessment event is constructed to employ all three dimensions in an integrated activity or set of activities (the pieces all come from the same places). Also, the assessment makes use of the vertical scaling functions of the learning progressions.	All three dimensions are addressed in combination across multiple assessment events.	At least two dimensions are addressed in balance, but primarily in isolation rather than integrated within the assessment task(s). Items are mapped to standards (and therefore maybe to ideas and practices) but actual items are not purposefully connected to each other.	Primarily addresses one dimension, or addresses additional dimensions minimally. Content is isolated in assessment (separate items with little integration).	The Iowa Core Science Standards categorize science into three integrated dimensions that need to all be addressed together in a balanced manner. The way the assessment is constructed should clearly show how the three dimensions are integrated. You can see how the assessment event* is constructed in a way that uses the science and engineering practices and/or the crosscutting practices to indicate the level of disciplinary core idea understanding (i.e., how deep is the knowledge). This is in direct contrast to a test built of a bunch of separate items, or at best, that has a few items centering around one idea, activity, or reading. * An "assessment event" is a series of related/integrated assessment tasks and performances that together allow for assessment of all three dimensions
Language is consistent with the Iowa Core Science Standards		Displays, reports and technical information match the organization and language of the three dimensions of the Iowa Core		Inconsistent, incomplete, or no match with terminology of the Iowa Core Science Standards	The intent of this item is to clarify whether language used in the assessment is in alignment with the Iowa Core Science Standards. Do the words that are used in the test documents match the words used in the Iowa Core? Can people use this without translation?

8. College/Career

1. Please demonstrate that your assessment measures progress toward college or career (content) readiness

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Ability to measure progress toward college or career (content) readiness		Statistical analysis measures progress toward college or career (content) readiness indicators.		No college or career statistical analysis available	Does the test offer a means to measure progress toward college or career (content) readiness? College/Career Content readiness relates directly to mastery of Iowa Core content (see also alignment rubric).

9. Technical Supports

Please demonstrate the technical supports that are available

1. Training on assessments and interpretation of reports
2. Availability of results - machine scored (including AI scored constructed response items)
3. Availability of results - human scored (student constructed responses)

Rubric item	Above Good	Good	So-So	Not OK	Comments
	3	2	1	0	
Training on assessments and interpretation of reports	Online modules for key elements, with instructional materials on administration and report interpretation (as in #2 and #1)	Instructional materials on administration and report interpretation (and manuals as described in #1)	Technical and user manuals with administration and interpretation information	Incomplete materials for administration and interpretation, no technical manual	Multiple training methods are desired, with flexible supports for key elements. Self-paced training options desired.
Availability of results - machine scored (including AI scored constructed response items)	Individual results available in real-time. Classroom, building and system available within 24 hours of last testing (receipt of student responses)	Group and individual reports available within one week of test completion (receipt of student responses)	Group and individual reports available within 2-3 weeks of receipt of student responses	Group and individual results available in more than 3 weeks of receipt of student responses	Level 3 assumes electronic assessment and reporting capability. Other levels allow for either electronic or paper administration and reporting.
Availability of results - human scored (student constructed responses)	Group and individual reports available within three weeks of receipt of student responses	Group and individual reports available within one month of receipt of student responses	Group and individual reports available within six weeks of receipt of student responses	Group and individual reports not available within six weeks of receipt of student	Used if human scoring required, split points between machine and human scoring.

Iowa Assessment Task Force RFI - Appendix A

Assessment Survey Part One

What follows are the specific information requests as defined by the Assessment Task Force in the accompanying Request for Information. Please use this form to organize the requested information for ease of review.

1. Fairness

Please demonstrate with a summary of practices and evidence, including any references to other documentation, how you assure fairness in your assessments.

1. Statistical evidence related to fairness in the development of the proposed assessment.

Enter summary here:
Enter any links to documents and page references here

2. Processes used to ensure fairness during test development (Universal Design)

Enter summary here:
Enter any links to documents and page references here

3. How the proposed assessment addresses fairness in administration through:
    a. Accommodations (accessibility)

Enter summary here:
Enter any links to documents and page references here

    b. Standardized directions

Enter summary here:
Enter any links to documents and page references here

c. Practice items

Enter summary here:
Enter any links to documents and page references here

2. Availability

Please summarize and reference any additional documentation about the availability of your assessments.

4. Grade availability (grades covered)

Enter summary here:
Enter any links to documents and page references here

5. Availability in last quarter of the school year

Enter summary here:
Enter any links to documents and page references here

3. Describes Achievement

Please provide summary and reference to any supporting documents to demonstrate how your assessments accurately describe student achievement.

6. Accurately describes student achievement

Enter summary here:
Enter any links to documents and page references here

7. Accurately describes growth

Enter summary here:

Enter any links to documents and page references here

8. Report availability

Enter summary here:
Enter any links to documents and page references here

4. Validity

Please provide summary and reference to any supporting documents to demonstrate how you have determined your assessment is valid

9. Criterion validity coefficients (correlational evidence)

Enter summary here:
Enter any links to documents and page references here

10. Description of methodology indicating the quality of validity evidence

Enter summary here:
Enter any links to documents and page references here

5. Reliability

Please provide summary and reference to any supporting documents to demonstrate how you have determined that your assessment is reliable

11. Internal consistency (alpha, split half, marginal)

Enter summary here:
Enter any links to documents and page references here

12. Stability over time (test retest, alternate form)

Enter summary here:
Enter any links to documents and page references here

13. Scorer consistency (inter-rater agreement in some form) (if applicable)

Enter summary here:
Enter any links to documents and page references here

14. Description of methodology indicating the quality of reliability evidence

Enter summary here:
Enter any links to documents and page references here

6. Piloted/Tested in Iowa

Please provide summary and reference to any supporting documents to demonstrate that your assessment has been piloted in Iowa

15. Piloted in Iowa (item tryout)

Enter summary here:
Enter any links to documents and page references here

16. Tested in Iowa (field tested)

Enter summary here:
Enter any links to documents and page references here

7. Alignment

Please provide summary and reference to any supporting documents to demonstrate that your assessment is aligned in the following ways:

17. Tables of specifications

Enter summary here:
Enter any links to documents and page references here

18. Coverage of disciplinary core ideas

Enter summary here:
Enter any links to documents and page references here

19. Coverage of scientific and engineering practices

Enter summary here:
Enter any links to documents and page references here

20. Coverage of crosscutting concepts

Enter summary here:
Enter any links to documents and page references here

21. Evidence of integration among the three dimensions of the Iowa Core Science Standards

Enter summary here:
Enter any links to documents and page references here

22. Language is consistent with the Iowa Core Science Standards

Enter summary here:
Enter any links to documents and page references here

8. College/Career

23. Please provide summary and reference to any supporting documents to demonstrate that your assessment measures progress toward college or career (content) readiness

Enter summary here:
Enter any links to documents and page references here

9. Technical Supports

Please provide summary and reference to any supporting documents to demonstrate the technical supports that are available

24. Training on assessments and interpretation of reports

Enter summary here:
Enter any links to documents and page references here

25. Availability of results - machine scored (including AI scored constructed response items)

Enter summary here:
Enter any links to documents and page references here

26. Availability of results - human scored (student constructed responses)

Enter summary here:
Enter any links to documents and page references here

Assessment Survey Part Two

The Assessment Task Force is interested in additional features of the proposed assessments, as listed below. However, responses to the questions below will not be scored in round one of the review. The Assessment Task Force will review and score responses to Part One of this survey, which will determine which vendors move into a second round of review.

The second round of review will focus on quality and accessibility of assessments and other issues related to implementation. The information below will be used during this second review. Additional information may be requested of assessment providers for the second round of assessment reviews.

1. Types of assessment items Part 1

Grade	% of total score from	Total number of	Comments
Grade 3			
Grade 4			
Grade 5			
Grade 6			
Grade 7			
Grade 8			
Grade 9			
Grade 10			
Grade 11			

2. Types of items Part 2

Grade	% of total score from Constructed Response (short answer response)	% of total score from Selected Response items	Comments
Grade 3			
Grade 4			
Grade 5			
Grade 6			
Grade 7			
Grade 8			
Grade 9			
Grade 10			
Grade 11			

3. What test format or delivery system is proposed? Check all that apply and/or list planned implementation dates if some proposed forms are still in development.

Content Area: Science				
Grade	**Paper Pencil Fixed Form**	**Computer-Based Fixed Form**	**Computer-Based Adaptive**	**Comments**
Grade 3				
Grade 4				
Grade 5				
Grade 6				
Grade 7				
Grade 8				
Grade 9				
Grade 10				
Grade 11				
Grade 12				

4. If computer-based and/or computer adaptive tests are proposed, what is the bridge plan for schools that currently do not have the technical capacity to administer them?

5. Describe the item types developed for each assessment form, including innovative item types.

6. How are paper/pencil test results equated with computer-based test results?

7. Which of the following did you use in test development?
   _____ Classical test theory
   _____ Item response theory
   _____ Generalizability theory
   _____ Other (please identify)

8. On which of the following devices and operating systems is the proposed assessment currently operating?
   _____ Windows 8 desktops/laptops
   _____ Windows 7 desktops/laptops
   _____ Mac OSX desktops/laptops
   _____ Chrome OS laptops
   _____ iOS tablets (iPads)
   _____ Android tablets
   _____ Windows 8 tablets
   _____ Other devices/operating systems:_____

9. What is the minimum bandwidth per student required for the proposed assessments?

   What is the recommended bandwidth per student?

10. What resources do you have to help schools plan for technology readiness? (e.g., help figuring out what is needed to buy or put in place)

11. What technical support do you offer to help schools as they schedule the administration of the assessments?

12. What technical support do you offer to help schools troubleshooting technical issues during the administration of the assessments (e.g., help desk)?

13. Please list all accommodations provided for the proposed assessments. Please note if accommodations are only available at certain grades, content areas, or administration formats.

    1. General accommodations and/or accessibility features (e.g., text readers)
    2. Accommodations for English Language Learners
    3. Accommodations for students with vision disabilities
    4. Accommodations for students with hearing impairments
    5. Accommodations for students with physical impairments
    6. Other specific accommodations

14. Do you offer a suite of aligned assessments related to the proposed summative assessment? If so, please list the other assessments in the suite and indicate their purpose. (i.e., formative, interim/benchmark, etc.)

15. List any other supporting resources for schools and teachers. (These could include practice tests, sample assessments, professional learning, etc.)

16. What is the estimated time to administer the proposed assessments? Indicate times or estimated times for each grade level and format if applicable.

Grade	Comment
Grade 3	
Grade 4	
Grade 5	
Grade 6	
Grade 7	
Grade 8	
Grade 9	
Grade 10	
Grade 11	

17. Please provide evidence of typical costs for the proposed assessment package. It is expected that this cost includes the following:

- A standard package of materials, supports, reports, and data files to deploy a summative assessment for Science for state accountability.
    - Data file with individual student data, including raw data, percentile rank, and scaled/standard scores per test (subtests, benchmarks, etc.) for transfer into data systems.
    - Printed and electronic reports by different groupings (i.e., disaggregation, but also by content divisions (domain, standard, etc.)
    - Materials and supports necessary for all of administration, scoring, and reporting, including recommended accommodations (i.e., braille, audio).

18. Please describe plans for revision or replacement of items or tests (longevity of assessment system).

19. Please describe how data privacy, data ownership, and data security are assured.

20. Please provide a list of at least five schools you have worked with in Iowa, including contact information.

21. Please confirm availability of the proposed assessment for operational implementation in the spring of 2017.

Appendix 3: Sample Performance Expectation

5-PS1 Matter and Its Interactions

5-PS1 Matter and Its Interactions

Students who demonstrate understanding can:

5-PS1-1. Develop a model to describe that matter is made of particles too small to be seen. [Clarification Statement: Examples of evidence supporting a model could include adding air to expand a basketball, compressing air in a syringe, dissolving sugar in water, and evaporating salt water.] [Assessment Boundary: Assessment does not include the atomic-scale mechanism of evaporation and condensation or defining the unseen particles.]

5-PS1-2. Measure and graph quantities to provide evidence that regardless of the type of change that occurs when heating, cooling, or mixing substances, the total weight of matter is conserved. [Clarification Statement: Examples of reactions or changes could include phase changes, dissolving, and mixing that form new substances.] [Assessment Boundary: Assessment does not include distinguishing mass and weight.]

5-PS1-3. Make observations and measurements to identify materials based on their properties. [Clarification Statement: Examples of materials to be identified could include baking soda and other powders, metals, minerals, and liquids. Examples of properties could include color, hardness, reflectivity, electrical conductivity, thermal conductivity, response to magnetic forces, and solubility; density is not intended as an identifiable property.] [Assessment Boundary: Assessment does not include density or distinguishing mass and weight.]

5-PS1-4. Conduct an investigation to determine whether the mixing of two or more substances results in new substances.

The performance expectations above were developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices

Developing and Using Models
Modeling in 3–5 builds on K–2 experiences and progresses to building and revising simple models and using models to represent events and design solutions.
- Develop a model to describe phenomena. (5-PS1-1)

Planning and Carrying Out Investigations
Planning and carrying out investigations to answer questions or test solutions to problems in 3–5 builds on K–2 experiences and progresses to include investigations that control variables and provide evidence to support explanations or design solutions.
- Conduct an investigation collaboratively to produce data to serve as the basis for evidence, using fair tests in which variables are controlled and the number of trials considered. (5-PS1-4)
- Make observations and measurements to produce data to serve as the basis for evidence for an explanation of a phenomenon. (5-PS1-3)

Using Mathematics and Computational Thinking
Mathematical and computational thinking in 3–5 builds on K–2 experiences and progresses to extending quantitative measurements to a variety of physical properties and using computation and mathematics to analyze data and compare alternative design solutions.
- Measure and graph quantities such as weight to address scientific and engineering questions and problems. (5-PS1-2)

Disciplinary Core Ideas

PS1.A: Structure and Properties of Matter
- Matter of any type can be subdivided into particles that are too small to see, but even then the matter still exists and can be detected by other means. A model showing that gases are made from matter particles that are too small to see and are moving freely around in space can explain many observations, including the inflation and shape of a balloon and the effects of air on larger particles or objects. (5-PS1-1)
- The amount (weight) of matter is conserved when it changes form, even in transitions in which it seems to vanish. (5-PS1-2)
- Measurements of a variety of properties can be used to identify materials. (Boundary: At this grade level, mass and weight are not distinguished, and no attempt is made to define the unseen particles or explain the atomic-scale mechanism of evaporation and condensation.) (5-PS1-3)

PS1.B: Chemical Reactions
- When two or more different substances are mixed, a new substance with different properties may be formed. (5-PS1-4)
- No matter what reaction or change in properties occurs, the total weight of the substances does not change. (Boundary: Mass and weight are not distinguished at this grade level.) (5-PS1-2)

Crosscutting Concepts

Cause and Effect
- Cause and effect relationships are routinely identified, tested, and used to explain change. (5-PS1-4)

Scale, Proportion, and Quantity
- Natural objects exist from the very small to the immensely large. (5-PS1-1)
- Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume. (5-PS1-2),(5-PS1-3)

Connections to Nature of Science

Scientific Knowledge Assumes an Order and Consistency in Natural Systems
- Science assumes consistent patterns in natural systems. (5-PS1-2)

Connections to other DCIs in fifth grade: N/A

Articulation of DCIs across grade-levels: **2.PS1.A** (5-PS1-1),(5-PS1-2),(5-PS1-3); **2.PS1.B** (5-PS1-2),(5-PS1-4); **MS.PS1.A** (5-PS1-1),(5-PS1-2),(5-PS1-3),(5-PS1-4); **MS.PS1.B** (5-PS1-2),(5-PS1-4)

Common Core State Standards Connections:

ELA/Literacy –

RI.5.7	Draw on information from multiple print or digital sources, demonstrating the ability to locate an answer to a question quickly or to solve a problem efficiently. *(5-PS1-1)*
W.5.7	Conduct short research projects that use several sources to build knowledge through investigation of different aspects of a topic. *(5-PS1-2),(5-PS1-3),*(5-PS1-4)
W.5.8	Recall relevant information from experiences or gather relevant information from print and digital sources; summarize or paraphrase information in notes and finished work, and provide a list of sources. *(5-PS1-2),(5-PS1-3),*(5-PS1-4)
W.5.9	Draw evidence from literary or informational texts to support analysis, reflection, and research. *(5-PS1-2),(5-PS1-3),(5-PS1-4)*

Mathematics –

MP.2	Reason abstractly and quantitatively. *(5-PS1-1),*(5-PS1-2),(5-PS1-3)
MP.4	Model with mathematics. *(5-PS1-1),*(5-PS1-2),(5-PS1-3)
MP.5	Use appropriate tools strategically. (5-PS1-2),(5-PS1-3)
5.NBT.A.1	Explain patterns in the number of zeros of the product when multiplying a number by powers of 10, and explain patterns in the placement of the decimal point when a decimal is multiplied or divided by a power of 10. Use whole-number exponents to denote powers of 10. *(5-PS1-1)*
5.NF.B.7	Apply and extend previous understandings of division to divide unit fractions by whole numbers and whole numbers by unit fractions. *(5-PS1-1)*
5.MD.A.1	Convert among different-sized standard measurement units within a given measurement system (e.g., convert 5 cm to 0.05 m), and use these conversions in solving multi-step, real-world problems. *(5-PS1-2)*
5.MD.C.3	Recognize volume as an attribute of solid figures and understand concepts of volume measurement. *(5-PS1-1)*
5.MD.C.4	Measure volumes by counting unit cubes, using cubic cm, cubic in, cubic ft, and improvised units. *(5-PS1-1)*

*The performance expectations marked with an asterisk integrate traditional science content with engineering through a Practice or Disciplinary Core Idea.

The section entitled "Disciplinary Core Ideas" is reproduced verbatim from A Framework for K-12 Science Education: Practices, Cross-Cutting Concepts, and Core Ideas. Integrated and reprinted with permission from the National Academy of Sciences.

Appendix 4: Depth-of-Knowledge Explanation and Examples

The following appendix is excerpted from Kentucky Department of Education, August 2007, *Support Materials for Core Content for Assessment Version 4.1 Science*, pages 2-9, accessed March 1, 2016, from https://www.aea267.k12.ia.us/system/assets/uploads/files/2472/science_samples.pdf.

Introduction to Depth of Knowledge (DOK) - Based on Norman Webb's Model (Karin Hess, Center for Assessment/NCIEA, 2005)

According to Norman L. Webb ("Depth of Knowledge Levels for Four Content Areas," March 28, 2002), interpreting and assigning depth of knowledge levels to both objectives within standards and assessment items is an essential requirement of alignment analysis.

Four Depth of Knowledge (DOK) levels were developed by Norman Webb as an alignment method to examine the consistency between the cognitive demands of standards and the cognitive demands of assessments.

Descriptors of DOK Levels for Science (based on Webb, March 2002)

Recall and Reproduction – Depth of Knowledge (DOK) Level 1
Recall and Reproduction requires recall of information, such as a fact, definition, term, or a simple procedure, as well as performing a simple science process or procedure. Level 1 only requires students to demonstrate a rote response, use a well-known formula, follow a set procedure (like a recipe), or perform a clearly defined series of steps. A "simple" procedure is well defined and typically involves only one-step. Verbs such as "identify," "recall," "recognize," "use," "calculate," and "measure" generally represent cognitive work at the recall and reproduction level. Simple word problems that can be directly translated into and solved by a formula are considered Level 1. Verbs such as "describe" and "explain" could be classified at different DOK levels, depending on the complexity of what is to be described and explained.

A student answering a Level 1 item either knows the answer or does not: that is, the answer does not need to be "figured out" or "solved." In other words, if the knowledge necessary to answer an item automatically provides the answer to the item, then the item is at Level 1. If the knowledge necessary to answer the item does not automatically provide the answer, the item is at least at Level 2.

Skills and Concepts/Basic Reasoning – Depth of Knowledge (DOK)

Level 2 Skills and Concepts/Basic Reasoning includes the engagement of some mental processing beyond recalling or reproducing a response. The content knowledge or process involved is more complex than in level 1. Items require students to make some decisions as to how to approach the question or problem. Keywords that generally distinguish a Level 2 item include "classify," "organize," "estimate," "make observations," "collect and display data," and "compare data." These actions imply more than one step. For example, to compare data requires first identifying characteristics of the objects or phenomenon and then grouping or ordering the objects. Level 2 activities include making observations and collecting data; classifying, organizing, and comparing data; and organizing and displaying data in tables, graphs, and charts.

Some action verbs, such as "explain," "describe," or "interpret," could be classified at different DOK levels, depending on the complexity of the action. For example, interpreting information from a simple graph, requiring reading information from the graph, is a Level 2. An item

that requires interpretation from a complex graph, such as making decisions regarding features of the graph that need to be considered and how information from the graph can be aggregated, is at Level 3.

Strategic Thinking/Complex Reasoning – Depth of Knowledge (DOK) Level 3

Strategic Thinking/Complex Reasoning requires deep knowledge using reasoning, planning, using evidence, and a higher level of thinking than the previous two levels. The cognitive demands at Level 3 are complex and abstract. The complexity does not result only from the fact that there could be multiple answers, a possibility for both Levels 1 and 2, but because the multi-step task requires more demanding reasoning. In most instances, requiring students to explain their thinking is at Level 3; requiring a very simple explanation or a word or two should be at Level 2. An activity that has more than one possible answer and requires students to justify the response they give would most likely be a Level 3. Experimental designs in Level 3 typically involve more than one dependent variable. Other Level 3 activities include drawing conclusions from observations; citing evidence and developing a logical argument for concepts; explaining phenomena in terms of concepts; and using concepts to solve non-routine problems.

Extended Thinking/Reasoning – Depth of Knowledge (DOK) Level 4

Extended Thinking/Reasoning requires high cognitive demand and is very complex. Students are required to make several connections—relate ideas within the content area or among content areas—and have to select or devise one approach among many alternatives on how the situation can be solved. Many on-demand assessment instruments will not include any assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated in such a way as to expect students to perform extended thinking. “Develop generalizations of the results obtained and the strategies used and apply them to new problem situations,” is an example of a Grade 8 objective that is a Level 4. Many, but not all, performance assessments and open-ended assessment activities requiring significant thought will be Level 4.

Level 4 requires complex reasoning, experimental design and planning, and **probably will require an extended period of time** either for the science investigation required by an objective, or for carrying out the multiple steps of an assessment item. However, the extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying significant conceptual understanding and higher-order thinking. For example, if a student has to take the water temperature from a river each day for a month and then construct a graph, this would be classified as a Level 2 activity. However, if the student conducts a river study that requires taking into consideration a number of variables, this would be a Level 4.

Table 1: Examples for each of the Depth of Knowledge Levels in Science, based on Webb

(Adapted from Karin Hess, Center for Assessment/NCIEA by the Kentucky Department of Education, 2005)

Recall & Reproduction (DOK 1)	Skills & Concepts/Basic Reasoning (DOK 2)	Strategic Thinking/ Complex Reasoning (DOK 3)	Extended Thinking/Reasoning (DOK 4)
a. Recall or recognize a fact,	a. Specify and explain the	a. Interpret information	a. Select or devise approach among

term, definition, simple procedure (such as one step), or property b. Demonstrate a rote response c. Use a well-known formula d. Represent in words or diagrams a scientific concept or relationship e. Provide or recognize a standard scientific representation for simple phenomenon f. Perform a routine procedure, such as measuring length g. Perform a simple science process or a set procedure (like a recipe) h. Perform a clearly defined set of steps i. Identify, calculate, or measure *NOTE: If the knowledge necessary to answer an item automatically provides the answer,*	relationship between facts, terms, properties, or variables b. Describe and explain examples and nonexamples of science concepts c. Select a procedure according to specified criteria and perform it d. Formulate a routine problem given data and conditions e. Organize, represent, and compare data f. Make a decision as to how to approach the problem g. Classify, organize, or estimate h. Compare data i. Make observations j. Interpret, information from a simple graph k. Collect and display data *NOTE: If the knowledge*	from a complex graph (such as determining features of the graph or aggregating data in the graph) b. Use reasoning, planning, and evidence c. Explain thinking (beyond a simple explanation or using only a word or two to respond) d. Justify a response e. Identify research questions and design investigations for a scientific problem f. Use concepts to solve non-routine problems/more than one possible answer g. Develop a scientific model for a complex situation h. Form conclusions from experimental or observational	many alternatives to solve problem b. Based on provided data from a complex experiment that is novel to the student, deduct the fundamental relationship between several controlled variables. c. Conduct an investigation, from specifying a problem to designing and carrying out an experiment, to analyzing its data and forming conclusions d. Relate ideas within the content area or among content areas e. Develop generalizations of the results obtained and the strategies used and apply them to new problem situations *NOTE: Level 4 activities often require an extended period of time for carrying out multiple steps; however, time*

it is a Level 1.	*necessary to answer an item does not automatically provide the answer, then the item is at least a Level 2. Most actions imply more than one step*	data i. Complete a multi-step problem that involves planning and reasoning j. Provide an explanation of a principle k. Justify a response when more than one answer is possible l. Cite evidence and develop a logical argument for concepts m. Conduct a designed investigation n. Research and explain a scientific concept o. Explain phenomena in terms of concepts	*alone is not a distinguishing factor if skills and concepts are simply repetitive over time.*

Depth of Knowledge as a "Ceiling"

Core Content statements are identified with a Depth of Knowledge (DOK) levels. This level represents the highest level (ceiling) that items will be designed for the Kentucky Core Content Test.

It is important to note, however, that items will also be developed below the ceiling level. Table 2 provides three examples of social studies core content statements with different "ceilings," that is, the highest DOK Level at which an item could be assessed. Table 2 also indicates the other DOK levels at which an item could be assessed.

Table 2: Examples of Science Assessment Standards and Potential Assessment Items

Sample Science Assessment Standard	Ceiling	Potential DOK Levels for Assessment
Example A: Perform a simple science process or a set procedure to gather data	1	1 (Measure temperature of water)
Example B: Represent data collected over a period time, making comparisons and interpretations 2	1	1 (Measure temperature of water at different times/places) 2 (Construct a graph to organize, display, and compare data)
Example C: Interpret data collected for a research question for a scientific problem related to your environment	3	1 (Measure temperature of water at different times/places) 2 (Construct a graph to organize, display, and compare data) 3 (Design an investigation to explain the effect of varying temperatures of the river in different locations) 3 (Design an investigation to explain the effect of varying temperatures of the river in different locations)

Table 3: Examples of Science Core Content Statements and Potential Assessment Items at all levels of DOK.

SC-07-1.1.1 Students will:

- **classify substances according to their chemical/reactive properties;**
- **infer real life applications for substances based on chemical/reactive properties.**

In chemical reactions, the total mass is conserved. Substances are often classified into groups if they react in similar ways. The patterns which allow classification can be used to infer or understand real life applications for those substances.

DOK 3

Recall & Reproduction	Skills & Concepts/Basic Reasoning	Strategic Thinking/Complex Reasoning	Extended Thinking/Reasoning
Describe the characteristics of an acid or base.	Create a chart classifying several different substances into established categories.	Create a chart listing practical applications of different substances and explain why those substances would be appropriate for the applications you suggest.	Conduct an investigation of different chemicals to determine which would be the best choice for a specific application, then test the results, analyze the data and form a conclusion.

SC-HS-4.6.8 Students will

- **describe the connections between the functioning of the Earth system and its sources of energy (internal and external).**
- **predict the consequences of changes to any component of the Earth system.**

Earth systems have sources of energy that are internal and external to the Earth. The Sun is the major external source of energy. Two primary sources of internal energy are the decay of radioactive isotopes and the gravitational energy from Earth's original formation.

DOK 3

Recall & Reproduction	Skills & Concepts/Basic Reasoning	Strategic Thinking/Complex Reasoning	Extended Thinking/Reasoning
Recite basic facts about Earth's energy sources. Recall definition of radioactive decay and gravitational energy. Do computations relating to half-life. Write a song or poem about Earth's energy sources.	Make observations. Collect and record data. Organize and display data in charts/tables. Explain the energy relationship between the Sun and the Earth Systems. Explain radioactive decay. Explain the role of	Generate a research question and design an investigation. Test effects of different variables. Student presentation of information (explain internal and external sources of energy). Defend a position (e.g., nuclear energy use, national energy policy).	Design and conduct an experiment. Student presentation of information (relate effect of human activity on availability of energy). Apply and adapt information to real world situations (e.g., C14 dating). Apply ideas outside of science context (e.g., economic impact of hybrid vehicles). Synthesize content

	gravity in Earth's formation.		from several resources (e.g., to make decisions regarding alternative power/fuel sources). Integrate concepts for a global understanding of energy (e.g., economics, environment, politics).

Depth of Knowledge is not verb dependent. It is based on the cognitive complexity of the standard.

Table 4: Same Verb used at 4 Depth of Knowledge Levels

VERB	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4
List	**List** three properties that can be used to classify objects.	**List** three groups that a wooden object could be classified as belonging in.	**List** three unique objects and the physical properties that would prevent each of the objects from sharing a category with any of the others.	**List** the design steps (including the controlled variables) you would take to investigate the best material for a specific purpose. Provide evidence from your investigation to support your conclusions.
Describe	**Describe** characteristics of metamorphic rocks.	**Describe** the difference between metamorphic and igneous rocks.	**Describe** a model that you might use to represent the relationships that exist within the rock cycle.	**Describe** the approach you would take to ensure that the rock samples you collect are truly representative of the geologic diversity of Kentucky.

Appendix 5: Resources

The following resources were used by the Assessment Task Force.

21st Century Science Assessment: The Future Is Now, https://www.sri.com/sites/default/files/publications/21st_century_science_assessment_pellegrino.pdf

A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas, National Academies Press, http://www.nap.edu/catalog/13165/a-framework-for-k-12-science-education-practices-crosscutting-concepts

ACT Aspire Exemplar Items, http://www.discoveractaspire.org/assessments/test-items/

ACT Aspire Periodic Assessments, http://www.discoveractaspire.org/assessments/periodic/

Developing Assessments for the Next Generation Science Standards, National Academies Press, http://www.nap.edu/catalog/18409/developing-assessments-for-the-next-generation-science-standards

Iowa Assessment Task Force Report, December 31, 2014, https://www.educateiowa.gov/sites/files/ed/documents/2014-12-31AssessmentTaskForceReport.pdf

Iowa Assessments Scope and Sequence, https://itp.education.uiowa.edu/ia/documents/Iowa_Form_E_F_Scope_and_Sequence.pdf

Iowa Core Science webpage, https://iowacore.gov/iowa-core/subject/science

Iowa Science Standards Implementation Plan, https://www.educateiowa.gov/sites/files/ed/documents/IowaScienceStandardsImplementationPlan_0.pdf

Next Generation Science Standards, http://www.nextgenscience.org/

Next Generation Science Standards Appendix A: Conceptual Shifts in the Next Generation Science Standards, http://www.nextgenscience.org/sites/default/files/Appendix%20A%20-%204.11.13%20Conceptual%20Shifts%20in%20the%20Next%20Generation%20Science%20Standards.pdf

Science Standards Review Team Report 2015, https://www.educateiowa.gov/sites/files/ed/documents/2015-08-06ScienceStandardsReviewTeamReport.pdf

Appendix 6: Iowa Code Section 256.7(21)(b)

256.7 Duties of the state board.

21. Develop and adopt rules incorporating accountability for, and reporting of, student achievement into the standards and accreditation process described in section 256.11. The rules shall provide for all of the following:
…
b. A set of core academic indicators in mathematics and reading in grades four, eight, and eleven, a set of core academic indicators in science in grades eight and eleven, and another set of core indicators that includes but is not limited to graduation rate, postsecondary education, and successful employment in Iowa.

(1) Annually, the department shall report state data for each indicator in the condition of education report. Rules adopted pursuant to this subsection shall specify that the approved district-wide assessment of student progress administered for purposes of the core academic indicators shall be the assessment utilized by school districts statewide in the school year beginning July 1, 2011, or a successor assessment administered by the same assessment provider.

(2) Notwithstanding subparagraph (1), for the school year beginning July 1, 2016, and each succeeding school year, the rules shall provide that all students enrolled in school districts in grades three through eleven shall be administered an assessment during the last quarter of the school year that at a minimum assesses the core academic indicators identified in this paragraph "*b*"; is aligned with the Iowa common core standards in both content and rigor; accurately describes student achievement and growth for purposes of the school, the school district, and state accountability systems; and provides valid, reliable, and fair measures of student progress toward college or career readiness.

(3) The director shall establish an assessment task force to review and make recommendations for a statewide assessment of student progress on the core academic indicators identified pursuant to this paragraph "*b*". The task force shall recommend a statewide assessment that is aligned to the Iowa common core standards and is, at a minimum, valid, reliable, tested, and piloted in Iowa. In addition, in developing recommendations, the task force shall consider the costs to school districts and the state in providing and administering such an assessment and the technical support necessary to implement the assessment. The task force shall submit its recommendations in a report to the director, the state board, and the general assembly by January 1, 2015. The task force shall assist with the final development and implementation of the assessment administered pursuant to subparagraph (2). The task force shall include but not be limited to teachers, school administrators, business leaders, representatives of state agencies, and members of the general public. This subparagraph is repealed July 1, 2020.

(4) The state board shall submit to the general assembly recommendations the state board deems appropriate for modifications of assessments of student progress administered for purposes of this paragraph "*b*".

Appendix 7: Task Force Meeting Schedule

Date	City	Facility	Time
September 30, 2015	Des Moines	Grimes Building	9:00 a.m. to 3:00 p.m.
November 10, 2015	Des Moines	Grimes Building	9:00 a.m. to 3:00 p.m.
December 15, 2015	West Des Moines	West Des Moines Learning Center	9:00 a.m. to 3:00 p.m.
January 20, 2016	West Des Moines	West Des Moines Learning Center	9:00 a.m. to 2:20 p.m.
February 16, 2016	Des Moines	Grimes Building	9:00 a.m. to 3:00 p.m.
March 11, 2016	Des Moines	virtual meeting	9:00 a.m. to 10:30 am and 2:30 p.m. to 3:30 p.m.

Subgroups and small groups also met:
October 27, 2015 in Cedar Rapids
October 30, 2015 in Des Moines
November 20, 2015, virtual meeting
January 15, 2016 in Cedar Rapids and Des Moines