Kansans CAN

Balanced Assessment System for Science

William Bennett
Center for Educational Testing and Evaluation

Lizette Burks
Kansas Department of Education

Objectives

- Identify a Balanced Assessment System for Science
- Analyze and Classify Examples of Science Tasks
- Recognize the Formative Assessment Process for Science
- Review Updates to the State Science Summative Assessment

Shrinking the Assessment Summative Footprint

For purposes of estimating assessment time, the Kansas State Department of Education calculates one session as 45 to 60 minutes.

	MATH Grades 3 - 8, 10			ELA (English Language Arts) Grades 3 - 8, 10			SCIENCE Grades 5, 8, 11	
	MACHINE SCORED	PERFORMANCE	TOTAL TIME *(minutes)*	MACHINE SCORED	PERFORMANCE	TOTAL TIME *(minutes)*	MACHINE SCORED	TOTAL TIME *(minutes)*
2016	4 sessions	1 session	225-300	4 sessions	2 sessions	270-360	3 sessions	135-180
2017	2 sessions	Moved to interim	90-120	2 sessions	Moved to interim	90-120	2 sessions	90-120

Source: Assessment Fact Sheet, July 2016, KSDE

Redirecting Focus

- Move away from over-emphasis on state summative standardized assessment
- Improve student learning through a richer instructionally focused system of assessment
- Renewed focus on the classroom, where the majority of assessment affecting students' daily lives occur
- Create a foundation of, and ongoing support for assessment literacy for all assessment users

Source: Assessment Literacy: Zach Conrad, August 2016, KSDE

	Type of Assessment		
	*"Assessment **for** Learning"*	*"Assessment **of** Learning"*	
Dimension	**Formative**	**Interim**	**Summative**
Purpose	▪ Instructional	▪ Most designed for managerial uses ▪ Some designed for instructional uses	▪ Managerial
Implementation	▪ Driven by moment-to-moment decisions; generated or selected by teacher; individualized	▪ Regulated by a set of rules developed in or out of the classroom; teacher-generated or externally generated	
Timing	▪ During instruction ▪ High frequency	▪ After instruction or during a break in instructional flow ▪ Moderate frequency	▪ After instruction ▪ Low frequency
Scope	▪ Narrow; one or very few learning objectives at a time	▪ Moderate; a manageable number of objectives	▪ Broad; comprehensive set of objectives
Audience	▪ Classroom (i.e., students, teachers, and parents)	▪ Administration and/or ▪ Classroom	▪ Public ▪ Administration ▪ Classroom
Feedback	▪ Student ⟷ teacher ▪ Descriptive/narrative	▪ School System ⟶ audiences ▪ Mostly evaluative	▪ School System ⟶ audiences ▪ Mostly evaluative

Talbot, T. (June 2011), *Comprehensive Assessment Systems: Purposes and Implementation*

Formative Chart

Assessment Type	Purpose	Frequency and Relationship to Instruction
Classroom Formative: Formal checkpoints on learning progress Classroom Formative: Embedded in ongoing teaching and learning (Student, Teacher)	Assist/evaluate teaching and learning Monitor learning relative to lesson goals Signal important learning goals Monitor progress with specifically targeted intervention	Minute-by-minute Daily Weekly During teaching and learning Or as fits with instructional plan or schedule

Source: Assessment Literacy: Zach Conrad, August 2016, KSDE

Interim Chart

Assessment Type	Purpose	Frequency and Relationship to Instruction
Classroom Summative Interim/bench mark Summative (Student, Teacher, SSP, Administrators, Family)	Motivate Signal important learning goals Evaluate achievement Monitor student learning, based on learning goals Predict end of year proficiency Inform improvement strategies for: • Teachers • Schools • Districts	After a more extended period of teaching and learning (e.g., after a unit is completed and before another unit begins) At the end of a semester 3x per year or more Across instructional units/calendar periods

Source: Assessment Literacy: Zach Conrad, August 2016, KSDE

Summative Chart

Assessment Type	Purpose	Frequency and Relationship to Instruction
Summative: State, district, school, other external mandated National & International Assessments (Student, Teacher, SSP, Administrators, Family, State, Policymakers, Public)	Signal important learning goals Accountability Identifying/prioritizing gross needs Informing/evaluating improvement strategies	After a year's worth or a course's worth of instruction and learning

Source: Assessment Literacy: Zach Conrad, August 2016, KSDE

Three Dimensional Instruction

Aligning Instruction in the Science Classroom

Performance Expectation →

Elements (Bullets) →

5-LS2-1 Ecosystems: Interactions, Energy, and Dynamics

Students who demonstrate understanding can:

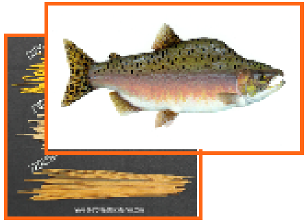
5-LS2-1. **Develop a model to describe the movement of matter among plants, animals, decomposers, and the environment.**[Clarification Statement: Emphasis is on the idea that matter that is not food (air, water, decomposed materials in soil) is changed by plants into matter that is food. Examples of systems could include organisms, ecosystems, and the Earth.] [Assessment Boundary: Assessment does not include molecular explanations.]

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

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Developing and Using Models Modeling in 3–5 builds on K–2 models and progresses to building and revising simple models and using models to represent events and design solutions. • Develop a model to describe phenomena. --------------------- ***Connections to the Nature of Science*** **Science Models, Laws, Mechanisms, and Theories Explain Natural Phenomena** • Science explanations describe the mechanisms for natural events.	**LS2.A: Interdependent Relationships in Ecosystems** • The food of almost any kind of animal can be traced back to plants. Organisms are related in food webs in which some animals eat plants for food and other animals eat the animals that eat plants. Some organisms, such as fungi and bacteria, break down dead organisms (both plants or plants parts and animals) and therefore operate as "decomposers." Decomposition eventually restores (recycles) some materials back to the soil. Organisms can survive only in environments in which their particular needs are met. A healthy ecosystem is one in which multiple species of different types are each able to meet their needs in a relatively stable web of life. Newly introduced species can damage the balance of an ecosystem. **LS2.B: Cycles of Matter and Energy Transfer in Ecosystems** • Matter cycles between the air and soil and among plants, animals, and microbes as these organisms live and die. Organisms obtain gases, and water, from the environment, and release waste matter (gas, liquid, or solid) back into the environment.	**Systems and System Models** • A system can be described in terms of its components and their interactions.

An Analogy between 3-Dimensional Learning and Cooking

Kitchen Tools & Techniques (Practices)



Basic Ingredients

(Core Ideas)

Vegetables, Herbs, Spices, & Seasonings

(Crosscutting Concepts)

Preparing a Meal

(Three dimensional Learning)

Example Science Tasks

Are Tasks 3 Dimensional?

5-LS2-1 Ecosystems: Interactions, Energy, and Dynamics

Students who demonstrate understanding can:

5-LS2-1. **Develop a model to describe the movement of matter among plants, animals, decomposers, and the environment.**[Clarification Statement: Emphasis is on the idea that matter that is not food (air, water, decomposed materials in soil) is changed by plants into matter that is food. Examples of systems could include organisms, ecosystems, and the Earth.] [Assessment Boundary: Assessment does not include molecular explanations.]

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

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Developing and Using Models Modeling in 3–5 builds on K–2 models and progresses to building and revising simple models and using models to represent events and design solutions. • Develop a model to describe phenomena. -------------------- ***Connections to the Nature of Science*** **Science Models, Laws, Mechanisms, and Theories Explain Natural Phenomena** • Science explanations describe the mechanisms for natural events.	**LS2.A: Interdependent Relationships in Ecosystems** • The food of almost any kind of animal can be traced back to plants. Organisms are related in food webs in which some animals eat plants for food and other animals eat the animals that eat plants. Some organisms, such as fungi and bacteria, break down dead organisms (both plants or plants parts and animals) and therefore operate as "decomposers." Decomposition eventually restores (recycles) some materials back to the soil. Organisms can survive only in environments in which their particular needs are met. A healthy ecosystem is one in which multiple species of different types are each able to meet their needs in a relatively stable web of life. Newly introduced species can damage the balance of an ecosystem. **LS2.B: Cycles of Matter and Energy Transfer in Ecosystems** • Matter cycles between the air and soil and among plants, animals, and microbes as these organisms live and die. Organisms obtain gases, and water, from the environment, and release waste matter (gas, liquid, or solid) back into the environment.	**Systems and System Models** • A system can be described in terms of its components and their interactions.

← Performance Expectation

← Elements (Bullets)

DCI

SEP

CCC

Example Science Tasks

Are the Tasks Formative, Interim, or Summative?

	Type of Assessment		
	*"Assessment **for** Learning"*	*"Assessment **of** Learning"*	
Dimension	**Formative**	**Interim**	**Summative**
Purpose	▪ Instructional	▪ Most designed for managerial uses ▪ Some designed for instructional uses	▪ Managerial
Implementation	▪ Driven by moment-to-moment decisions; generated or selected by teacher; individualized	▪ Regulated by a set of rules developed in or out of the classroom; teacher-generated or externally generated	
Timing	▪ During instruction ▪ High frequency	▪ After instruction or during a break in instructional flow ▪ Moderate frequency	▪ After instruction ▪ Low frequency
Scope	▪ Narrow; one or very few learning objectives at a time	▪ Moderate; a manageable number of objectives	▪ Broad; comprehensive set of objectives
Audience	▪ Classroom (i.e., students, teachers, and parents)	▪ Administration and/or ▪ Classroom	▪ Public ▪ Administration ▪ Classroom
Feedback	▪ Student ⟷ teacher ▪ Descriptive/narrative	▪ School System ⟶ audiences ▪ Mostly evaluative	▪ School System ⟶ audiences ▪ Mostly evaluative

Talbot, T. (June 2011), *Comprehensive Assessment Systems: Purposes and Implementation*

Formative Assessment Process

FORMATIVE ASSESSMENT PLANNING GUIDE

THE CENTER ON STANDARDS & ASSESSMENT IMPLEMENTATION
WestEd CRESST

LEARNING GOALS

What is the learning intended by the end of this lesson?

SUCCESS CRITERIA

What will students do to show they are progressing toward the Learning Goal?

EVIDENCE-GATHERING OPPORTUNITIES

How will teachers and students collect information about students' progress toward the Learning Goal?

TASKS, RESOURCES, STEPS & STRATEGIES

What else do teachers need to plan in order to implement this lesson?

Revisit Example Science Tasks

Are the Tasks Formative, Interim, or Summative?

	Type of Assessment		
	*"Assessment **for** Learning"*	*"Assessment **of** Learning"*	
Dimension	**Formative**	**Interim**	**Summative**
Purpose	▪ Instructional	▪ Most designed for managerial uses ▪ Some designed for instructional uses	▪ Managerial
Implementation	▪ Driven by moment-to-moment decisions; generated or selected by teacher; individualized	▪ Regulated by a set of rules developed in or out of the classroom; teacher-generated or externally generated	
Timing	▪ During instruction ▪ High frequency	▪ After instruction or during a break in instructional flow ▪ Moderate frequency	▪ After instruction ▪ Low frequency
Scope	▪ Narrow; one or very few learning objectives at a time	▪ Moderate; a manageable number of objectives	▪ Broad; comprehensive set of objectives
Audience	▪ Classroom (i.e., students, teachers, and parents)	▪ Administration and/or ▪ Classroom	▪ Public ▪ Administration ▪ Classroom
Feedback	▪ Student ⟷ teacher ▪ Descriptive/narrative	▪ School System ⟶ audiences ▪ Mostly evaluative	▪ School System ⟶ audiences ▪ Mostly evaluative

Talbot, T. (June 2011), *Comprehensive Assessment Systems: Purposes and Implementation*

State Science Summative Assessment Update

- 2016 Field test
- Differences between 2016 and 2017
- 2017 Operational test notes
- Embedded field testing process
- Released items and examples
- Questions and answers

Transition to KCSS

Field test of newly developed items for the new standards

- Purpose is to analyze new items for their statistical performance
  - Difficulty, discrimination, and reliability
  - Items with poor reliability and discrimination are discarded or revised
  - Note: no items were flagged for bias

Overview of field test format

- No items from previous 2015 were included
  - Performance between standards could not be equated
- Length of assessment
  - Grade 5: 60 items
  - Grades 8 and 11: 70 items

Overview of Differences

2016 Field Test & 2017 Operational Test

Following the 2016 field test, several changes were made to the test design for the 2017 operational form

- Length of assessment reduced to shorten assessment footprint and difficulty
- Number of stages also reduced
- Embedded field test items will contain a new media type, simulations
- General population students can now use a limited text-to-speech (TTX) option

Features staying constant

- Assessment sessions will remain two 45-60 minutes periods
- Content emphasis will preserve last year’s depth and breadth of coverage

Reduction in Assessment Footprint

Item reduction and testing time

2016 Grade 5 field test contained 60 items per student

- Approximately 1.5-2 minutes per item

2017 Grade 5 operational test will contain 39 items per student

- About 35% reduction overall
- Approximately 2-3 minutes per item

2016 Grade 8 and 11 field tests contained 70 items per student

- Approximately 1.5 minutes per item

2017 Grade 8 and 11 operational tests will contain 44 items per student

- About 37% reduction overall
- Approximately 2-3 minutes per item

All grades will have items organized into 2 stages with 1 stage per session

Embedded Field Testing for 2017

New computer simulation performance tasks will be field tested

- Chemistry and Ecology

Computer simulations are interactive rich media

- Allow students to observe results or data from the choices they make
- Provides better opportunities to assess student performances over non-interactive items
- Machine-scored responses avoids problems associated with hand-scoring
- Adaptable to each grade level
- Expandable over time to allow for new and different tasks and observations

Simulation Formats

Both simulations use sets of 4 items that share a single media

- Chemistry simulation items
  - Students can drag-and-drop reactants into a test tube to observe changes in properties
  - Students use the observations and data to answer associated items
- Ecology simulation items
  - Students use buttons to select different starting population values of different organisms to interact
  - Observations of interaction animations and graphs of populations are used as data to answer associated items

Continued Development of Science Items

Need for new items determined by statistics and test direction (ex. new items for use in simulation sets)

- Content experts develop rough drafts
- Rough drafts get first edit
- Teacher review
- Revision of items
  - Content, 3-D alignment, text length, clarity, bias, accessibility, etc.
- Revised items get another edit
- Field testing
  - Statistical performance of item
- Review of assessment to begin next round of development

Content Emphasis

- Past distribution of domains is maintained in shorter assessments
- Not all items can be evenly distributed among Claims and Targets
  - All areas of the standards are important
  - Some areas have different numbers than others
    - Item specifications are on the ksassessment.org website
    - Ex. 5^{th} grade Structure and Properties of Matter ~22-27%
    - Differences arise due grade-level differences and field-test items

Content by Grade Overview

Grade 5

- Claim 1, physical science: 35-40%
- Claim 2, life science: 24-29%
- Claim 3, earth and space science: 33-38%

Grade 8

- Claim 1, physical science: 34-38%
- Claim 2, life science: 31-36%
- Claim 3, earth and space science: 26-32%

Grade 11

- Claim 1, physical science: 27-33%
- Claim 2, life science: 34-40%
- Claim 3, earth and space science: 27-33%

Released Item Grade 5

Ruby conducts an experiment to determine the best conditions to grow radishes from seeds. She arranges four small cups on a windowsill. The medium and water conditions she used for each plant are shown in this table.

Cup	Medium	Water
1	Shredded paper towels	Yes
2	Soil	Yes
3	Soil	No
4	Shredded paper towels	No

Which results would support the claim that growing plants need certain resources to grow?

- Radishes start to grow in cups 2 and 3 but not in cups 1 and 4 because the radishes need soil to grow. (23%)
- Radishes start to grow in all four cups because the radishes receive enough sunlight to grow. (10%)
- Radishes start to grow in cups 1 and 2 but not in cups 3 and 4 because the radishes need water to grow. (56%)
- Radishes will not grow in any of the four cups because the radishes need to be outside in a garden to grow. (10%)

Connections to Standards and Extensions

Claim 2, Target A, Level 3: “Students can … determine requirements of plant growth from observations…”

- 5-LS1-1: Support an argument that plants get the materials they need for growth chiefly from air and water.

If the same type of item were to be used at a lower level/complexity

- identify the components or factors that are used as evidence
- For example, identify air or water

If the same type of item were to be used a higher level/complexity

- analyze data in an item to use as evidence to support a claim
- Data table listing soil and plant weight over time

Released Item Grade 8

Jacob had a plant with red flowers. He exposed seeds from that plant to a chemical. He grows two plants, plant 1 and plant 2, from those seeds. Plant 1 produced red flowers, and plant 2 produced white flowers. Jacob sequenced each plant's DNA and found that both plants had mutations. When he compared the protein sequences from each plant, plant 1 had a red pigmentation protein that was the same length as the parent's protein, while plant 2 produced a red pigmentation protein that was only half as long as the other two plants' proteins. Why was the color of plant 2's flowers different?

- The red color comes from proteins determined by gene sequences, which were damaged by the chemical. (41%)
- The chemical absorbed the pigments from the flowers, making them white. (21%)
- The chemical damaged the proteins, which made the flowers white. (23%)
- The red color comes from pigments in chloroplasts, which are invisible as long as green chlorophyll is produced. (13%)

Connections to Standards and Extensions

Claim 2, Target D, Level 3: "Students can… distinguish differences in organism traits to differences in genes and reproduction."

- MS-LS3-1 - Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins...

If this same type of item were used at a lower level/complexity

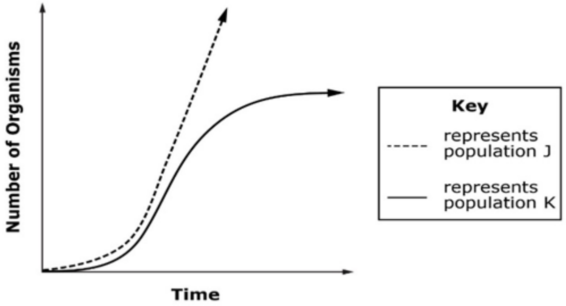
- recognize that genes are affected by reproduction and mutation
- attribute the cause of changes in protein structure to changes in gene structure (DNA)

If this same type of item were used at a high level/complexity

- use a model of a gene or DNA as part of an explanation for changes to protein structure
- describe of how a mutation changes a diagram model of a protein made from a normal gene

Released Item Grade 11

This graph illustrates the growth of two organism populations, J and K, over time.



Which factor may have caused population K to have a lower growth over time compared to population J?

- increase in competition for food (47%)
- increase of genetic diversity (15%)
- high birthrate of individuals (15%)
- short lifespan of individuals (20%)

Connections to Standards and Extensions

"Students can connect changes in living and nonliving factors to population growth…"

- HS-LS2-1 - Use mathematical and/or computational representations to support explanations of factors that affect carrying capacity of ecosystems…

If this same type of item were asked at a lower level/complexity

- identify the resources required by populations to grow over time
- select a variable (ex. competition) which causes populations to decline over time

If this same type of item were asked at a higher level/complexity

- explain the interaction that influence carrying capacity in a mathematical expression
- explain why a graph showing population growth over time is lower/higher for a population with two factors interacting instead of only one factor

Resources

Formative Lesson Planning Guide Template:

- http://tinyurl.com/zccftcp

Creating 3-Dimensional Tasks:

- http://stemteachingtools.org/brief/30

MS 3-Dimensional Task Project:

- http://nextgenscienceassessment.org

KS Assessments (CETE):

- http://ksassessments.org

Kansans CAN

Balanced Assessment System for Science

William Bennett: Email- william-d-bennett@ku.edu
Center for Educational Testing and Evaluation

Lizette Burks: Email- lburks@ksde.org, Twitter- @lizette_burks, #KSSci
Kansas Department of Education