Council of Chief State School Officers

Science Assessment Item Collaborative

Grade 5 Item Cluster Prototype

for assessment of the

Next Generation Science Standards

November 2015

Developed by WestEd in collaboration with CCSSO Science Assessment Item Collaborative state members and content experts.

Introduction

On behalf of the Council of Chief State School Officers (CCSSO) Science Assessment Item Collaborative (SAIC) state members, WestEd's science content and assessment specialists are pleased to present this Grade 5 Item Cluster Prototype for assessment of the Next Generation Science Standards (NGSS). This prototype was designed to follow the principles and recommendations set forth in the SAIC Assessment Framework and Item Specifications Guidelines for an NGSS-aligned large-scale summative assessment item cluster. This first iteration of the item cluster prototype has undergone many rounds of refinement, and its developers are optimistic that, if it is considered in conjunction with key sections of the Assessment Framework, it will serve as an initial building block for the development of large-scale summative assessments measuring the NGSS.

Development of the Grade 5 Item Cluster Prototype was a collaborative effort that depended upon the significant expertise and experience of SAIC state members, science content experts, and assessment designers and developers. It is anticipated that future iterations of item clusters will be informed by this foundational work and by future work from the wider science research and assessment community. The developers of this prototype are hopeful that this future work continues to reflect the goal of providing states and other jurisdictions with a vetted approach to ensuring that emerging NGSS-based assessments are fair, meet the highest standards for technical quality, and are aligned to the principles of responsible testing articulated in the report *Developing Assessments for the Next Generation Science Standards* (hereafter "the BOTA report") (NRC, 2014).

Intended Uses of This Prototype

1. To serve as an initial model for measuring the three-dimensional science learning called for in the NGSS (NGSS Lead States, 2013) and in *A Framework for K–12 Science Education* (hereafter "the *K–12 Framework*") (NRC, 2012).
2. To support states in guiding their test vendors with the design and development of NGSS-aligned assessments. *This may be accomplished by presenting the prototype as a model that can be customized for each state's unique context.*

   Issues of feasibility must be considered by the state and by the vendor, as these issues will affect implementation. These issues include platform capability, practicality, cost, time to develop, time to administer, number of constructed-response (CR) items, and scoring methods.
3. To promote ongoing dialogue, in the science education and assessment communities, about the vision for a truly next-generation science assessment and the opportunity to develop a research-supported, innovative, large-scale assessment for measuring the NGSS.

The following sections outline the working assumptions under which the Grade 5 Item Cluster Prototype was developed and provide a summary of the decisions and accompanying rationales for features of the prototype that warranted additional discussion during development.

Prototype Assumptions

1. The prototype was developed with the intent to show how the Assessment Framework could be implemented for the purpose of development of an NGSS-aligned large-scale summative assessment. As such, the Assessment Framework, the *K–12 Framework*, the BOTA report, the NGSS, and the NGSS Evidence Statements (NGSS Network, 2015a, 2015b) are at the foundation of the prototype. Those seeking to use this approach will benefit from deep familiarity with these resources.
2. The prototype was developed with input from SAIC state members. The prototype development process did not require full consensus on every issue, but does represent the best thinking of state members.
3. The Assessment Framework was written prior to development of the prototype. Thus, some minor discrepancies emerged between the prototype and recommendations in the Assessment Framework.
4. The prototype is presented as a series of static item cards, not as functional items. This is indicative of the SAIC's focus on alignment rather than on system functionality. The annotations and metadata provided in the item cards are necessary to understand the intent of the items and the overall scaffolding in the item cluster.
5. The prototype was designed to be delivery system–agnostic. General principles of computer-based assessment delivery systems were used in developing the prototype, but the item structure and functionality described in the prototype are not meant to be representative of any specific delivery system. User interface (UI) notes are presented to help guide the interpretation of the intended system functionality.

Design Decisions, Execution, and Process Summary

Template Structure. The overall design of an item card includes student view (left), UI notes (right), and alignment information and item part metadata (at the bottom).

In addition, stimulus slides consist of the student view (left) and UI notes (right). Item overview slides are provided to show the complete presentation of each item to the student, inclusive of all item parts but without UI notes.

Alignment. The item cluster, when taken in its entirety, is intended to achieve three-dimensional alignment to the targeted NGSS Performance Expectations (PEs). As stated in the NGSS Evidence Statements Executive Summary (NGSS Network, 2015a),

> *Each PE represents the integration of three "dimensions" of science education: scientific and engineering practices (SEPs), disciplinary core ideas (DCIs), and crosscutting concepts (CCCs). As such, both student learning and assessment around the NGSS should be "three dimensional."* (p. 1)

Items, inclusive of all their parts, are intended to align at the PE and dimension (SEP, DCI, and CCC) levels.[1] Discrete item parts are aligned to one or more Evidence Statements associated with the target PEs. This organization is intended to directly support overall item alignment to PEs and dimensions. Information on this level of alignment is provided on the item card pages and in the comprehensive metadata table included in the prototype. As stated in the NGSS Evidence Statements Executive Summary (NGSS Network, 2015a),

> *The evidence statements are meant to show what it looks like for students to fully satisfy the PE [and] . . . were designed to articulate how students can use the practices to demonstrate their understanding of the DCIs through the lens of the CCCs, and thus, demonstrate proficiency on each PE.* (p. 1)

Scaffolding. The Grade 5 Item Cluster Prototype is designed to assess students along a range of proficiency and across an appropriate range of cognitive complexity. The intent of scaffolding in this context is to provide a structure that allows students of all abilities to demonstrate their proficiency through progressively built item parts and purposeful sequencing of items. The NGSS Evidence Statements were written to describe student evidence at a single proficiency level and to cover a range of cognitive complexity. In order to serve the needs of a large-scale summative assessment, the item cluster provides scaffolding with respect to anticipated item difficulty, independent from alignment to the sequence of the Evidence Statements, and in order to effectively collect a range of information along a proficiency continuum.

Items and Item Parts. Most items within the item cluster prototype have multiple parts. This allows for more complex interactions and deeper thinking—and allows for the employment of science practices on the part of the student. The presumed navigational control that is offered by online administration is leveraged in order to scaffold the items within the item cluster. Students may navigate freely to the stimulus at any time (preferably through a tabbed structure) and can navigate freely between parts within an item. Students cannot navigate back to items that have been previously submitted. For example, students can change their responses to Part (a) and Part (b) of a single item at will, but cannot change their response to a previous item once the item is submitted and the student has navigated to the next item in the sequence.

Item Types and Scoring Considerations. SAIC member states believe that the effort and time needed to score CR items must be considered when selecting item types for inclusion in a cluster. In order to balance this consideration with the overall alignment goals, every effort was made to select items of appropriate types—i.e., those item types that offered the ideal functionality for the measurement purpose—while limiting the number of likely hand-scored items (e.g., CR items). While CR items can be an effective means of measuring complex three-dimensional learning, technology-enhanced items (TEIs), evidence-based selected-response items (EBSRs), and selected-response items (SRs) were included when they were deemed effective at achieving alignment and serving as a valid measure of the intended constructs.

Content-Related Decisions. During the development process, certain choices were deemed necessary by the SAIC in order to satisfy the majority of its members. For example, one such choice is the use of "mass," instead of exclusively using "weight," for PE 5-PS1-2. The PE and DCI boundaries state that "Mass and weight are not distinguished at this grade and should not be assessed." After much deliberation, it was decided that the term "mass" would be used in all provided information, as it is more scientifically accurate, but that "mass" and "weight" would be accepted equally in scoring student responses. It is expected that states may choose to modify this specific approach to fit their individual needs.

Use of Prototype to Emphasize Various Design Elements. The SAIC members wanted a range of Assessment Framework design elements to be reflected in the prototypes that emerge from their collaboration. These ranges of design elements will be evident in two significant ways: (1) via stimulus innovation, interaction, and item dependency, and (2) via the relative number of CR items in the item cluster. These ranges reflect the intended utility of the prototypes.

[1] See the Assessment Framework and the appendix of this prototype for a more in-depth description of item cluster alignment expectations. Additional information also can be found on the Item Overview slide pages and in the comprehensive metadata table that immediately follows the last item in the item cluster.

Level:	Grade 5
Primary Target Domain:	Physical Sciences
Target PEs:	5-PS1-1, 5-PS1-2
Crosscutting Concept(s) Focus:	Scale, Proportion, and Quantity
Science and Engineering Practice(s) Focus:	Developing and Using Models, Using Mathematics and Computational Thinking
Reasoning for PE Groupings:	Mass (size micro to macro), and conservation of mass
Phenomenon:	Sugar is no longer visible when it dissolves in water, but the mass of the mixture stays the same
Allowable Item Types:	SR, TE, CR

	5-PS1-1	**5-PS1-2**
Performance Expectations:	Develop a model to describe that matter is made of particles too small to be seen.	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.
Target Clarifications:	Examples of evidence could include adding air to expand a basketball, compressing air in a syringe, dissolving sugar in water, and evaporating salt water.	Examples of reactions or changes could include phase changes, dissolving, and mixing that form new substances.
Assessment Boundary:	Assessment does not include the atomic-scale mechanism of evaporation and condensation or defining the unseen particles.	Assessment does not include distinguishing mass and weight.

Disciplinary Core Idea(s):	**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.	**PS1.A: Structure and Properties of Matter** • The amount (weight) of matter is conserved when it changes form, even in transitions in which it seems to vanish. **PS1.B: Chemical Reactions** • 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.)
Science and Engineering Practice(s):	**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. • Use models to describe phenomena.	**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.
Crosscutting Concept(s):	**Scale, Proportion, and Quantity** • Natural objects exist from the very small to the immensely large.	**Scale, Proportion, and Quantity** • Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume. *Connections to Nature of Science* **Scientific Knowledge Assumes an Order and Consistency in Natural Systems** • Science assumes consistent patterns in natural systems.

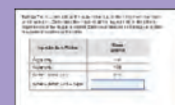
5-PS1-1 Develop a model to describe that matter is made of particles too small to be seen.

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.

ITEM CLUSTER

Stimulus

Item 1: Text Entry / Table Fill-In — **5-PS1-2** SEP CCC



Item 2a: Computation

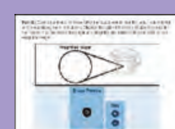
Item 2b: Graphing (TEI)

Item 2c: Short Answer

5-PS1-2 DCI SEP CCC

Item 3(a–b): Multiple Choice / Multiple Select — **5-PS1-2** DCI SEP CCC

Item 4a: Short Answer



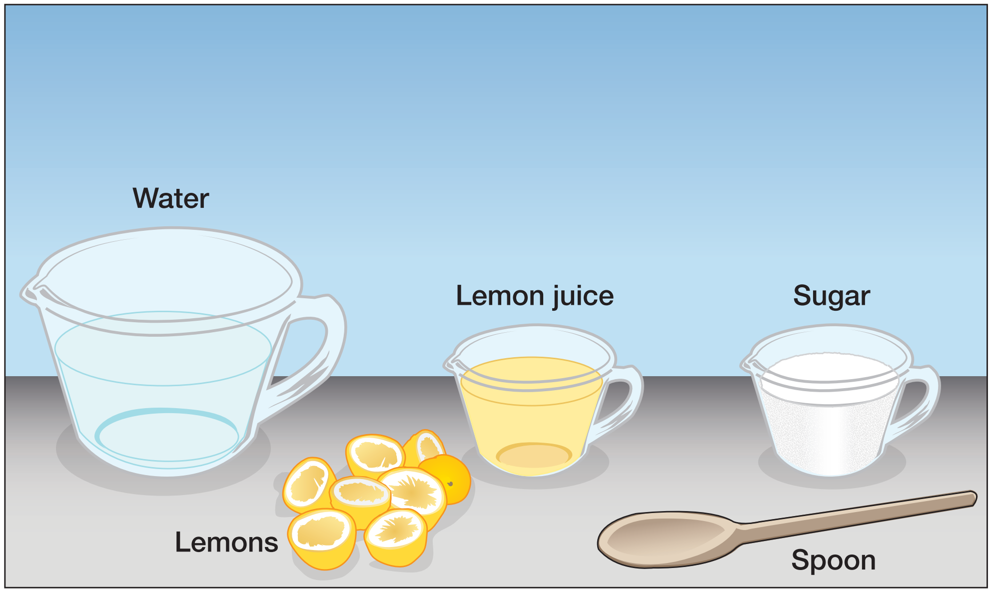
Item 4b: Building a Model (Drag-and-Drop, TEI)

5-PS1-1 DCI SEP CCC

Item 5(a–b): Constructed Response — **5-PS1-1** DCI SEP CCC

Students are making lemonade using three ingredients: water, lemon juice, and sugar. While making the lemonade, the students plan to investigate how the ingredients are affected by a physical change that happens when the ingredients are stirred together. The students measure out how much of each ingredient they need by following a recipe.

Stimulus



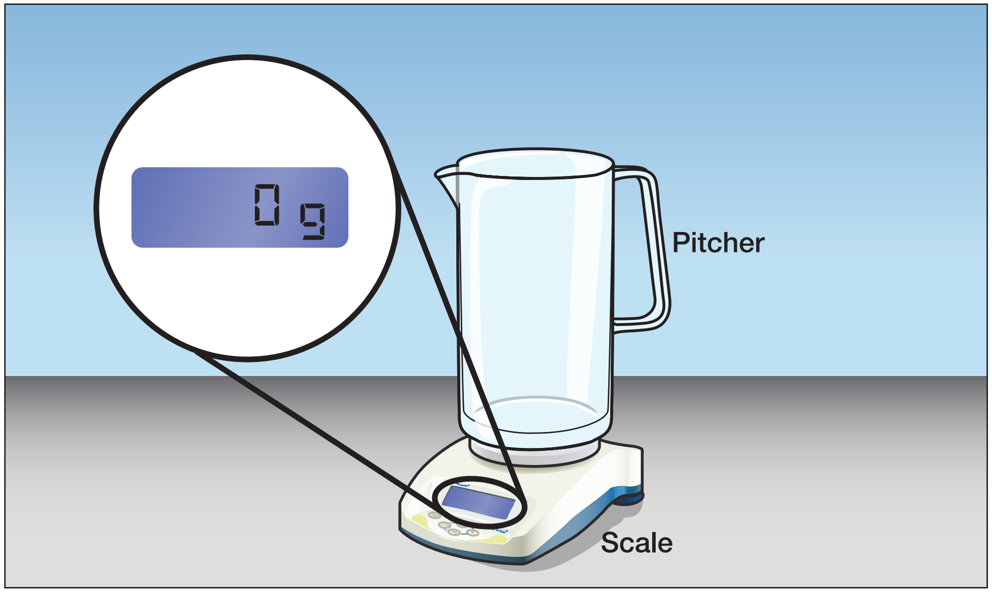
Media (animation/video): Stimulus is introduced; zoom in to each ingredient as it is presented to the student. It is recommended that the video include an audio narration.

REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue NEXT

Platform prompt and student control

The students use a digital scale in their investigation. The students will make the lemonade in a pitcher, but the students do not want to include the pitcher in the measurement on the scale. The students put the pitcher on the scale and reset the scale so it reads 0 g before adding any ingredients to the pitcher.



REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue. NEXT

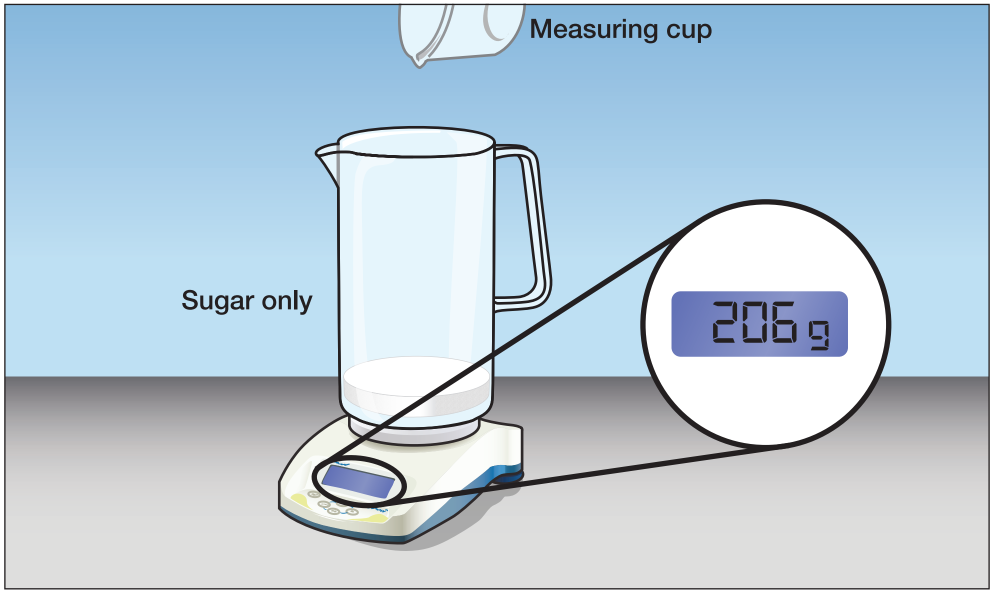
Stimulus

Media (animation/video): At the start, the scale reads 0 g to indicate that the scale has been tared.

The blowup of the reading on the scale is for illustrative purposes only. The final media will allow students to view/access the reading at any time without cluing the subsequent items.

Platform prompt and student control

The students pour the sugar from the measuring cup into the pitcher and record the information shown on the scale.



REPLAY Click REPLAY to watch the animation/video again.

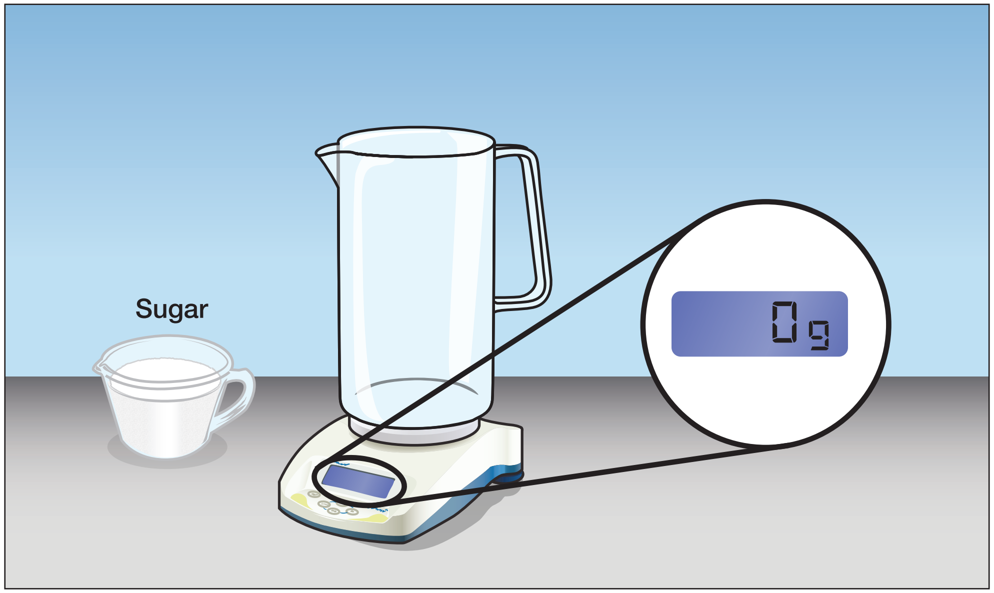
Click NEXT to continue. NEXT

Stimulus

Media (animation/video): All of the sugar is poured into the pitcher and the scale reads 206 g. Mass reading on the scale changes gradually as the sugar is being added to the pitcher.

Platform prompt and student control

The students then pour the sugar back into the measuring cup after the information is recorded.



REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue. NEXT

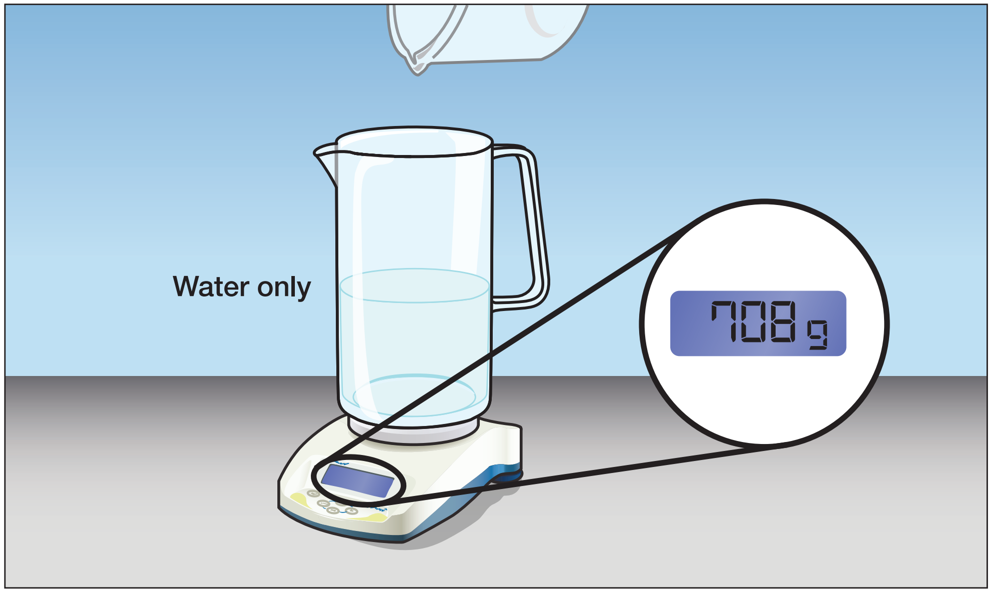
Stimulus

Media (animation/video): All of the sugar is poured into the measuring cup and the scale reads 0 g.

Platform prompt and student control

Next, the students pour water into the empty pitcher on the scale.

Stimulus



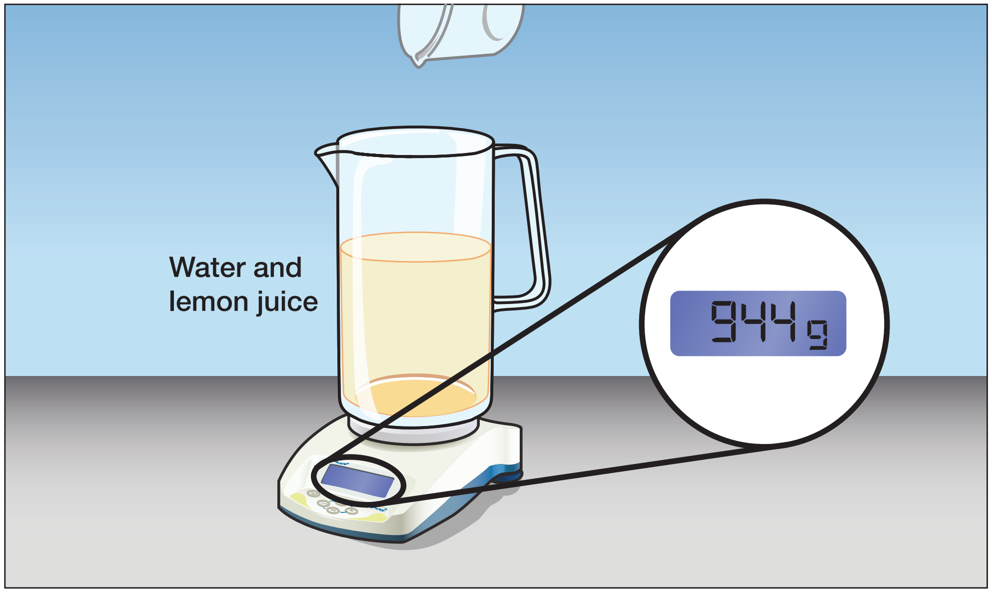
Media (animation/video): All of the water is poured into the pitcher and the scale reads 708 g. Mass reading on the scale changes gradually as an ingredient is being added.

REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue. NEXT

Platform prompt and student control

Then, the students pour the lemon juice into the pitcher of water, as shown.



Stimulus

Media (animation/video): All of the lemon juice is poured into the pitcher and the scale reads 944 g. Mass reading on the scale changes gradually as an ingredient is being added.

REPLAY Click REPLAY to watch the animation/video again.

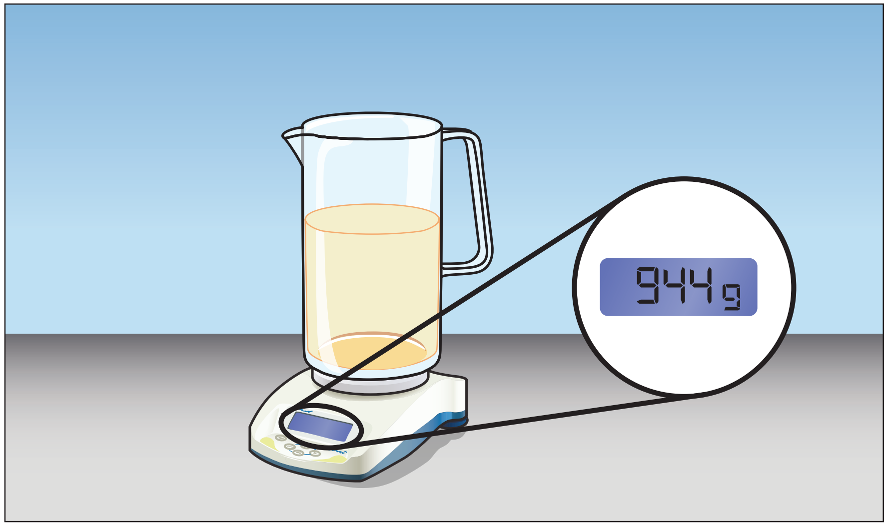
Click NEXT to continue. NEXT

Platform prompt and student control

The students want to determine how much of each ingredient will be in the lemonade. **Replay the animation by clicking the "REPLAY" button at the bottom of the screen. Type in the correct labels for the type of measurement you are collecting and then complete the table with the data you collect while replaying the animation.**

Ingredients	Measurement: ___ Units: ___
Sugar only	
Water only	
Water + lemon juice	

ANIMATION / VIDEO:



REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue to the next question. NEXT

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.

Full alignment to the PE and targeted dimensions is intended through the entirety of the item cluster. Partial to strong alignment to the dimensions for each item is achieved through alignment to the evidence statements, and is inclusive of all item parts for any given item.

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.

Scale, Proportion, and Quantity
- Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume.

The students want to determine how much of each ingredient will be in the lemonade. **Replay the animation by clicking the "REPLAY" button at the bottom of the screen. Type in the correct labels for the type of measurement you are collecting and then complete the table with the data you collect while replaying the animation.**

Ingredients	Measurement: ____ Units: ____
Sugar only	
Water only	
Water + lemon juice	

REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue to the next question. NEXT

Stem

Student enters one response in each of the text entry fields. If student enters only "g" in the "Units" field and clicks NEXT, a prompt appears instructing students to enter the full name of the unit.

Student populates the table while viewing the animation/video. There is a pause after each ingredient is added in order to allow the student to record the final mass for each ingredient in the data table. Entries in the table should be restricted to numeric only, with an appropriate character limit. (The student can also pause manually.)

Platform prompt and student control: The "REPLAY" button changes to a "PAUSE" button after the animation/video begins playing, and the student can toggle between "PLAY" and "PAUSE" to the end of the animation/video. Upon completion of the animation/video, the "REPLAY" button again appears.

Item Type: Text Entry / Table Fill-In
Estimated Time: 2 min

Evidence Statement Alignment:
(5-PS1-2)
(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (i) The weight of substances before they are heated, cooled, or mixed.

(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (ii) The weight of substances, including any new substances produced by a reaction, after they are heated, cooled, or mixed.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student can measure the mass of the lemonade mixture (or ingredients) in standard units (grams) before and after the mixture is made. Note that students will graph these quantities in Item 2. The student provides the labels in the table in order to achieve alignment to the CCC.

The students want to determine how much of each ingredient will be in the lemonade. **Replay the animation by clicking the "REPLAY" button at the bottom of the screen. Type in the correct labels for the type of measurement you are collecting and then complete the table with the data you collect while replaying the animation.**

Ingredients	Measurement: **Mass** Units: **grams**
Sugar only	**206**
Water only	**708**
Water + lemon juice	**944**

REPLAY Click REPLAY to watch the animation/video again.

Click NEXT to continue to the next question. NEXT

Stem

Student entered one response in each of the text entry fields. If student enters only "g" in the "Units" field and clicks NEXT, a prompt appears instructing students to enter the full name of the unit.

Student populated table

Platform prompt and student control: The "REPLAY" button changes to a "PAUSE" button after the animation/video begins playing, and the student can toggle between "PLAY" and "PAUSE" to the end of the animation/video. Upon completion of the animation/video, the "REPLAY" button again appears.

Item Type: Text Entry / Table Fill-In
Estimated Time: 2 min

Evidence Statement Alignment:
(5-PS1-2)
(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (i) The weight of substances before they are heated, cooled, or mixed.

(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (ii) The weight of substances, including any new substances produced by a reaction, after they are heated, cooled, or mixed.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student can measure the mass of the lemonade mixture (or ingredients) in standard units (grams) before and after the mixture is made. Note that students will graph these quantities in Item 2. The student provides the labels in the table in order to achieve alignment to the CCC.

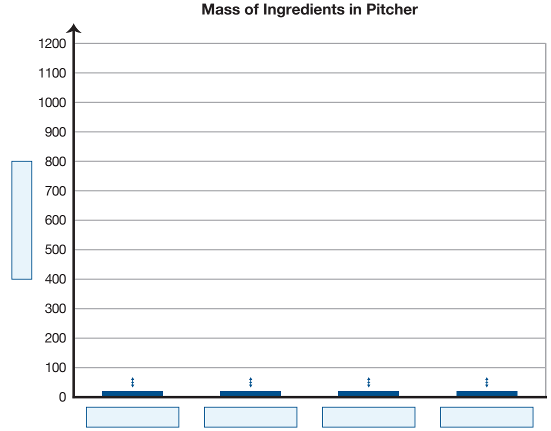
Scoring Notes: 2 points – 1 point is awarded for the correct data entered into the table; 1 point is awarded for the correct labels for the type of measurement and units. Use of the term "mass" is to reflect scientific accuracy. Students are not expected to differentiate mass from weight. Students will receive full credit for entering either "Mass" or "Weight" for the type of measurement.

Part (a) The students add all the sugar in the cup to the pitcher with the water and lemon juice. **Determine the total mass of all the ingredients in the pitcher once the sugar is added. Enter your answer, including units, into the correct location in the table.**

Ingredients in Pitcher	Mass (grams)
Sugar only	206
Water only	708
Water + lemon juice	944
Water + lemon juice + sugar	

Part (b) Now you will graph the data you collected. **Complete the graph to show the mass of the ingredients in the pitcher after each ingredient is added. Click on the top of the bar to drag and change the height of each bar. Then, type in a label in the appropriate space below each bar. Type in the appropriate label along the vertical axis (be sure to include an appropriate unit).**

Mass of Ingredients in Pitcher



Part (c) After stirring, the students observe that none of the sugar could be seen in the lemonade mixture. **Explain how the mass of the ingredients in the pitcher right after the sugar is added compares to the mass of the ingredients after the sugar is stirred.**

Click NEXT to continue to the next question.

NEXT

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.

Full alignment to the PE and targeted dimensions is intended through the entirety of the item cluster. Partial to strong alignment to the dimensions for each item is achieved through alignment to the evidence statements, and is inclusive of all item parts for any given item.

PS1.A: Structure and Properties of Matter
- The amount (weight) of matter is conserved when it changes form, even in transitions in which it seems to vanish.

PS1.B: Chemical Reactions
- 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.)

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.

Scale, Proportion, and Quantity
- Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume.

Part (a) The students add all the sugar in the cup to the pitcher with the water and lemon juice. **Determine the total mass of all the ingredients in the pitcher once the sugar is added. Enter your answer, including units, into the correct location in the table.**

Ingredients in Pitcher	Mass (grams)
Sugar only	206
Water only	708
Water + lemon juice	944
Water + lemon juice + sugar	

Click NEXT to continue to the next question.

NEXT

Stem

*Item 2 Parts (a), (b), and (c) will appear together on the same screen, and students may change their responses to Part (a), Part (b), or Part (c) at their discretion before clicking NEXT and continuing to Item 3. Students may **not** return to Item 1 at this stage in the administration.*

Student enters an alphanumeric response in this field. It is recommended that students have access to a calculator for this item. A character limit will be included for the response field.

Platform prompt and student control

Item Type: Computation
Estimated Time: 1 min

Evidence Statement Alignment:

(5-PS1-2)

(2) Mathematical/computational analysis: (a) Students measure and/or calculate the difference between the total weight of the substances (using standard units) before and after they are heated, cooled, and/or mixed.

(2) Mathematical/computational analysis: (c) Students use their measurements and calculations to describe that the total weights of the substances did not change, regardless of the reaction or changes in properties that were observed.

Note on Item Alignment:

What is being elicited from the student (evidence)? The student can calculate the mass of the sugar added to the liquid and reason that the mass of the sugar did not change when it was added to the liquid, even though the sugar was no longer visible in the liquid (i.e., it dissolved).

Part (a) The students add all the sugar in the cup to the pitcher with the water and lemon juice. **Determine the total mass of all the ingredients in the pitcher once the sugar is added. Enter your answer, including units, into the correct location in the table.**

Stem

Ingredients in Pitcher	Mass (grams)
Sugar only	206
Water only	708
Water + lemon juice	944
Water + lemon juice + sugar	**1150**

Item 2 Parts (a), (b), and (c) will appear together on the same screen, and students may change their responses to Part (a), Part (b), or Part (c) at their discretion before clicking NEXT and continuing to Item 3. Students may ***not*** *return to Item 1 at this stage in the administration.*

Student entered an alphanumeric response in this field.

Click NEXT to continue to the next question.

NEXT

Platform prompt and student control

Item Type: Computation
Estimated Time: 1 min

Evidence Statement Alignment:
(5-PS1-2)
(2) Mathematical/computational analysis: (a) Students measure and/or calculate the difference between the total weight of the substances (using standard units) before and after they are heated, cooled, and/or mixed.

(2) Mathematical/computational analysis: (c) Students use their measurements and calculations to describe that the total weights of the substances did not change, regardless of the reaction or changes in properties that were observed.

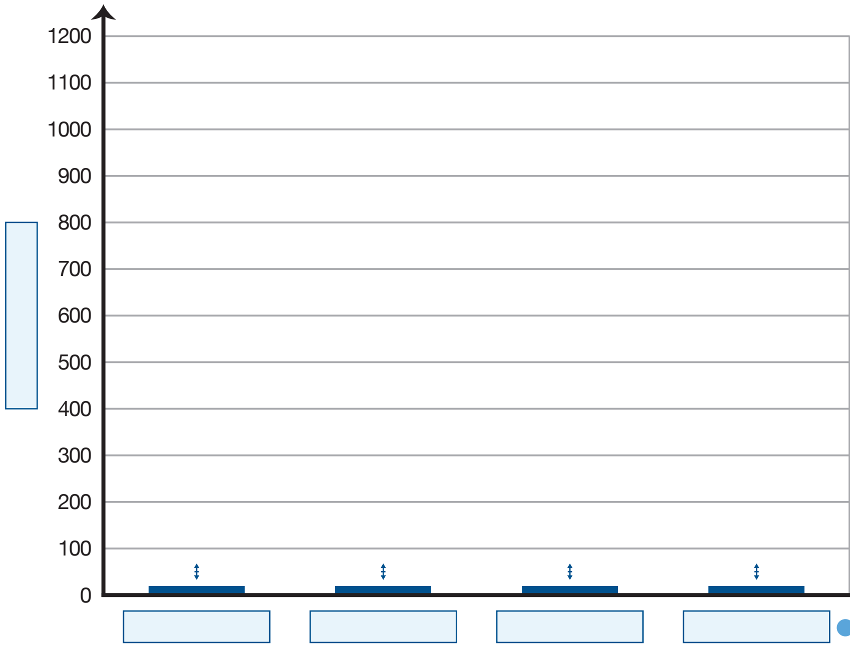
Note on Item Alignment:
What is being elicited from the student (evidence)? The student can calculate the mass of the sugar added to the liquid and reason that the mass of the sugar did not change when it was added to the liquid, even though the sugar was no longer visible in the liquid (i.e., it dissolved).

Scoring Notes: 1 point is awarded for the correct alphanumeric response.

Part (b) Now you will graph the data you collected. **Complete the graph to show the mass of the ingredients in the pitcher after each ingredient is added. Click on the top of the bar to drag and change the height of each bar. Then, type in a label in the appropriate space below each bar. Type in the appropriate label along the vertical axis (be sure to include an appropriate unit).**

Stem

Mass of Ingredients in Pitcher



Student clicks on the grid above a category to create a vertical bar showing the mass of that category. The student can click again to adjust the bars.

Student enters graph labels in these fields. Fields expand as needed.

Click NEXT to continue to the next question.

NEXT

Platform prompt and student control

Stimulus viewable by the student

Ingredients in Pitcher	Mass (grams)
Sugar only	206
Water only	708
Water + lemon juice	944
Water + lemon juice + sugar	1150

*Item 2 Parts (a), (b), and (c) will appear together on the same screen, and students may change their responses to Part (a), Part (b), or Part (c) at their discretion before clicking NEXT and continuing to Item 3. Students may **not** return to Item 1 at this stage in the administration.*

Item Type: Graphing
Estimated Time: 2 min

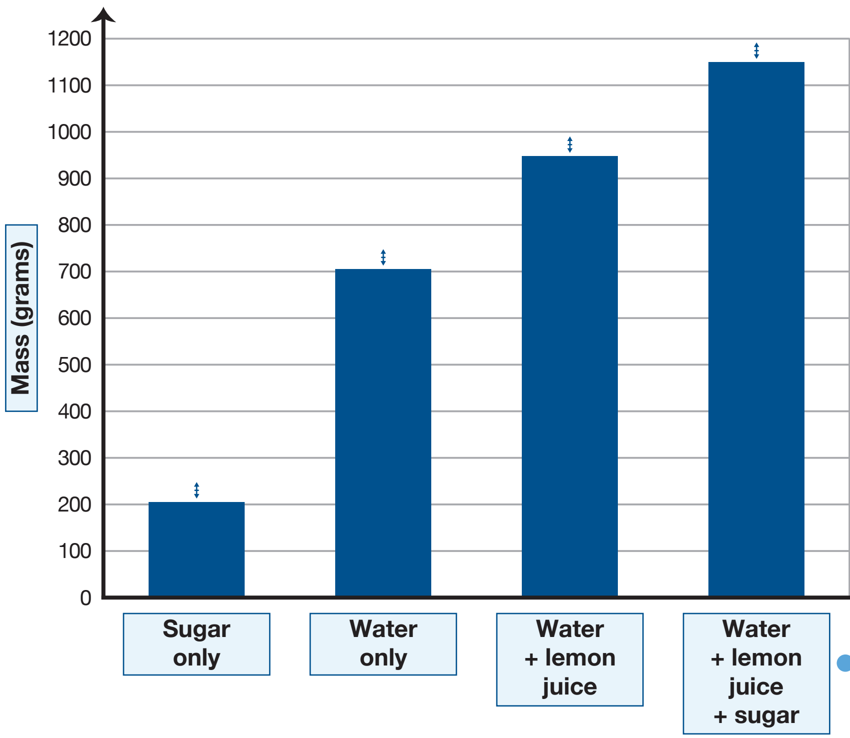
Evidence Statement Alignment:
(5-PS1-2)
(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (i) The weight of substances before they are heated, cooled, or mixed.

(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (ii) The weight of substances, including any new substances produced by a reaction, after they are heated, cooled, or mixed.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student can graph the masses of the lemonade mixture (or ingredients) before and after the mixture is made. The student graphs the data he or she collected, but is not penalized if the data was incorrect as long as the graph is correct in order for the two items to distinguish between the student's ability to measure from the student's ability to graph.

Part (b) Now you will graph the data you collected. **Complete the graph to show the mass of the ingredients in the pitcher after each ingredient is added. Click on the top of the bar to drag and change the height of each bar. Then, type in a label in the appropriate space below each bar. Type in the appropriate label along the vertical axis (be sure to include an appropriate unit).**

Mass of Ingredients in Pitcher



Click NEXT to continue to the next question.

NEXT

Stem

Stimulus viewable by the student

Ingredients in Pitcher	Mass (grams)
Sugar only	206
Water only	708
Water + lemon juice	944
Water + lemon juice + sugar	1150

Item 2 Parts (a), (b), and (c) will appear together on the same screen, and students may change their responses to Part (a), Part (b), or Part (c) at their discretion before clicking NEXT and continuing to Item 3. Students may ***not*** *return to Item 1 at this stage in the administration.*

Student created a vertical bar graph showing the mass of each category.

Student entered graph labels in these fields.

Platform prompt and student control

Item Type: Graphing
Estimated Time: 2 min

Evidence Statement Alignment:
(5-PS1-2)
(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (i) The weight of substances before they are heated, cooled, or mixed.

(1) Representation: (a) Students measure and graph the given quantities using standard units, including: (ii) The weight of substances, including any new substances produced by a reaction, after they are heated, cooled, or mixed.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student can graph the masses of the lemonade mixture (or ingredients) before and after the mixture is made. The student graphs the data he or she collected, but is not penalized if the data was incorrect as long as the graph is correct in order for the two items to distinguish between the student's ability to measure from the student's ability to graph.

Scoring Notes: 2 points are awarded for four correct bars with correct labels, and 1 point is awarded for 2–3 correct bars with correct labels, with a pre-established level of precision (recommended within +/-10 g).

Part (c) After stirring, the students observe that none of the sugar could be seen in the lemonade mixture. **Explain how the mass of the ingredients in the pitcher right after the sugar is added compares to the mass of the ingredients after the sugar is stirred.**

Click NEXT to continue to the next question.

NEXT

Stem

Student types responses in this field. Field expands as needed.

Platform prompt and student control

Stimulus viewable by the student

Ingredients in Pitcher	Mass (grams)
Sugar only	206
Water only	708
Water + lemon juice	944
Water + lemon juice + sugar	1150

*Item 2 Parts (a), (b), and (c) will appear together on the same screen, and students may change their responses to Part (a), Part (b), or Part (c) at their discretion before clicking NEXT and continuing to Item 3. Students may **not** return to Item 1 at this stage in the administration.*

Key

The mass of the ingredients is the same before and after the sugar is stirred even though it can no longer be seen.

Item Type: Short Answer
Estimated Time: 2 min

Evidence Statement Alignment:

(5-PS1-2)

(2) Mathematical/computational analysis: (a) Students measure and/or calculate the difference between the total weight of the substances (using standard units) before and after they are heated, cooled, and/or mixed.

(2) Mathematical/computational analysis: (c) Students use their measurements and calculations to describe that the total weights of the substances did not change, regardless of the reaction or changes in properties that were observed.

Note on Item Alignment:

What is being elicited from the student (evidence)? The student can reason that the weight of the sugar did not change after stirring even though it was no longer visible in the lemonade mixture.

Scoring Notes: 1 point is awarded for a correct response; see Key.

Part (a) The sugar could not be seen after the mixture was stirred. Which statement **best** explains what happened to the sugar?

- ○ The sugar was destroyed by the liquids.
- ○ The sugar became liquid water when stirred.
- ○ The sugar separated into particles too small to be seen.
- ○ The sugar was changed into a new substance by the lemon juice.

Part (b) Based on both the experiment presented here and your knowledge, which statement(s) below provide strong evidence to support your answer to Part (a)? Select all that apply.

- ○ The mixture was stirred after the sugar was added.
- ○ The mass of the mixture did not change after stirring.
- ○ The ingredients added to the mixture were almost all liquids.
- ○ The mass of the mixture increased when the sugar was added.
- ○ The sugar looked the same before and after it was added to the mixture.

Click NEXT to continue to the next question.

NEXT

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.

Full alignment to the PE and targeted dimensions is intended through the entirety of the item cluster. Partial to strong alignment to the dimensions for each item is achieved through alignment to the evidence statements, and is inclusive of all item parts for any given item.

PS1.A: Structure and Properties of Matter
- The amount (weight) of matter is conserved when it changes form, even in transitions in which it seems to vanish.

PS1.B: Chemical Reactions
- 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.)

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.

Scale, Proportion, and Quantity
- Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume.

Part (a) The sugar could not be seen after the mixture was stirred. Which statement **best** explains what happened to the sugar?

- ◯ The sugar was destroyed by the liquids.
- ◯ The sugar became liquid water when stirred.
- ◯ The sugar separated into particles too small to be seen.
- ◯ The sugar was changed into a new substance by the lemon juice.

Click NEXT to continue to the next question.

NEXT

Stem

Student selects one response from the choices.

Item 3 Parts (a) and (b) will appear together on the same screen, and students may change their responses to Part (a) or Part (b) at their discretion before clicking NEXT and continuing to Item 4. Students may ***not*** *return to Items 1 or 2 at this stage in the administration.*

Platform prompt and student control

Item Type: Multiple Choice
Estimated Time: 2 min *(for Parts a and b)*

Evidence Statement Alignment:
(5-PS1-2)
(2) Mathematical/computational analysis: (d) Students use measurements and descriptions of weight, as well as the assumption of consistent patterns in natural systems, to describe evidence to address scientific questions about the conservation of the amount of matter, including the idea that the total weight of matter is conserved after heating, cooling, or mixing substances.

Note on Item Alignment:
What is being elicited from the student (evidence)? In combination with Item 3 Part (b), the student can use the data from the experiment (measurements) and his or her understanding of consistent patterns in nature to provide evidence that the total weight of matter is conserved during the experiment. Distractors were formulated based on a common misconception that students hold regarding properties of matter (http://www.rsc.org/images/Misconceptions_update_tcm18-188603.pdf): "Matter has no permanent aspect. When matter disappears from sight (e.g., when sugar dissolves in water) it ceases to exist."

Part (a) The sugar could not be seen after the mixture was stirred. Which statement **best** explains what happened to the sugar?

- ○ The sugar was destroyed by the liquids.
- ○ The sugar became liquid water when stirred.
- ● The sugar separated into particles too small to be seen.
- ○ The sugar was changed into a new substance by the lemon juice.

Click NEXT to continue to the next question.

NEXT

Stem

Student selected one response from the choices.

Item 3 Parts (a) and (b) will appear together on the same screen, and students may change their responses to Part (a) or Part (b) at their discretion before clicking NEXT and continuing to Item 4. Students may ***not*** *return to Items 1 or 2 at this stage in the administration.*

Platform prompt and student control

Item Type: Multiple Choice
Estimated Time: 2 min *(for Parts a and b)*

Evidence Statement Alignment:
(5-PS1-2)
(2) Mathematical/computational analysis: (d) Students use measurements and descriptions of weight, as well as the assumption of consistent patterns in natural systems, to describe evidence to address scientific questions about the conservation of the amount of matter, including the idea that the total weight of matter is conserved after heating, cooling, or mixing substances.

Note on Item Alignment:
What is being elicited from the student (evidence)? In combination with Item 3 Part (b), the student can use the data from the experiment (measurements) and his or her understanding of consistent patterns in nature to provide evidence that the total weight of matter is conserved during the experiment. Distractors were formulated based on a common misconception that students hold regarding properties of matter (http://www.rsc.org/images/Misconceptions_update_tcm18-188603.pdf): "Matter has no permanent aspect. When matter disappears from sight (e.g., when sugar dissolves in water) it ceases to exist."

Scoring Notes: 1 point is awarded for the correct response to Part (a).

Part (b) Based on both the experiment presented here and your knowledge, which statement(s) below provide strong evidence to support your answer to Part (a)? Select all that apply.

- ○ The mixture was stirred after the sugar was added.
- ○ The mass of the mixture did not change after stirring.
- ○ The ingredients added to the mixture were almost all liquids.
- ○ The mass of the mixture increased when the sugar was added.
- ○ The sugar looked the same before and after it was added to the mixture.

Click NEXT to continue to the next question. NEXT

Stem

Student selects from the choices.

Students will have the option to replay the video in the initial stimulus at any point during the item cluster.

*Item 3 Parts (a) and (b) will appear together on the same screen, and students may change their responses to Part (a) or Part (b) at their discretion before clicking NEXT and continuing to Item 4. Students may **not** return to Items 1 or 2 at this stage in the administration.*

Platform prompt and student control

Item Type: Multiple Select
Estimated Time: 2 min *(for Parts a and b)*

Evidence Statement Alignment:
(5-PS1-2)
(2) Mathematical/computational analysis: (d) Students use measurements and descriptions of weight, as well as the assumption of consistent patterns in natural systems, to describe evidence to address scientific questions about the conservation of the amount of matter, including the idea that the total weight of matter is conserved after heating, cooling, or mixing substances.

Note on Item Alignment:
What is being elicited from the student (evidence)? In combination with Item 3 Part (a), the student can use the data from the experiment (measurements) and his or her understanding of consistent patterns in nature to provide evidence that the total weight of matter is conserved during the experiment. Distractors were formulated based on common misconceptions that students hold regarding properties of matter (http://www.rsc.org/images/Misconceptions_update_tcm18-188603.pdf): “Matter has a materialistic core to which various random properties having independent existence are attached. Matter can ‘disappear,’ whereas its properties (such as sweetness) can continue to exist completely independently of it.” “Weight is not an intrinsic property of matter. The existence of weightless matter can be accepted.”

Part (b) Based on both the experiment presented here and your knowledge, which statement(s) below provide strong evidence to support your answer to Part (a)? Select all that apply.

- ○ The mixture was stirred after the sugar was added.
- ● The mass of the mixture did not change after stirring.
- ○ The ingredients added to the mixture were almost all liquids.
- ● The mass of the mixture increased when the sugar was added.
- ○ The sugar looked the same before and after it was added to the mixture.

Click NEXT to continue to the next question. NEXT

Stem

Student selected two responses from the choices.

*Item 3 Parts (a) and (b) will appear together on the same screen, and students may change their responses to Part (a) or Part (b) at their discretion before clicking NEXT and continuing to Item 4. Students may **not** return to Items 1 or 2 at this stage in the administration.*

Platform prompt and student control

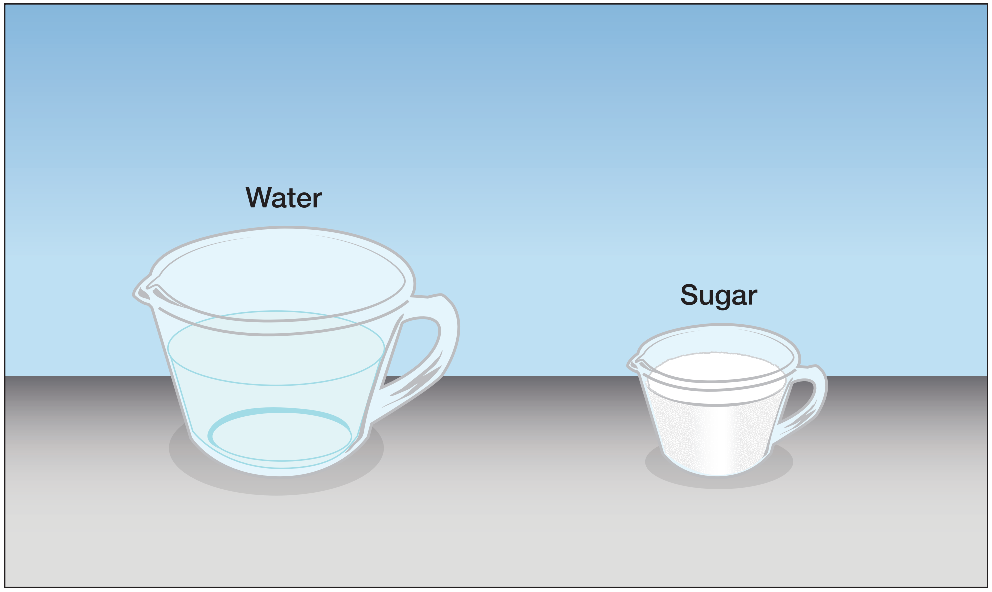
Item Type: Multiple Select
Estimated Time: 2 min *(for Parts a and b)*

Evidence Statement Alignment:
(5-PS1-2)
(2) Mathematical/computational analysis: (d) Students use measurements and descriptions of weight, as well as the assumption of consistent patterns in natural systems, to describe evidence to address scientific questions about the conservation of the amount of matter, including the idea that the total weight of matter is conserved after heating, cooling, or mixing substances.

Note on Item Alignment:
What is being elicited from the student (evidence)? In combination with Item 3 Part (a), the student can use the data from the experiment (measurements) and his or her understanding of consistent patterns in nature to provide evidence that the total weight of matter is conserved during the experiment. Distractors were formulated based on common misconceptions that students hold regarding properties of matter (http://www.rsc.org/images/Misconceptions_update_tcm18-188603.pdf): “Matter has a materialistic core to which various random properties having independent existence are attached. Matter can ‘disappear,’ whereas its properties (such as sweetness) can continue to exist completely independently of it.” “Weight is not an intrinsic property of matter. The existence of weightless matter can be accepted.”

Scoring Notes: 1 point is awarded for selecting the two correct responses to Part (b).

The students gather more water and sugar to investigate each of these ingredients more closely.



Click NEXT to continue. NEXT

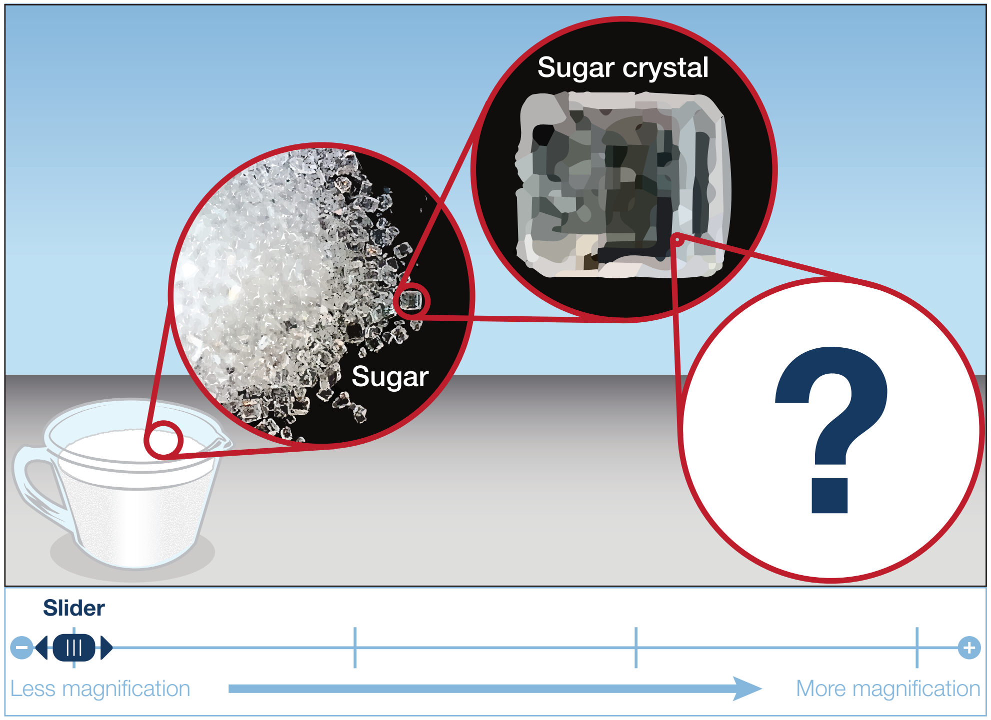
Stimulus

Graphic

Platform prompt and student control

The students look closely at the sugar and water and then research to find images of what the ingredients look like when magnified using a powerful microscope.

Slide the slider to the right and left to observe what sugar looks like at different magnifications.



Click NEXT to continue. NEXT

Stimulus

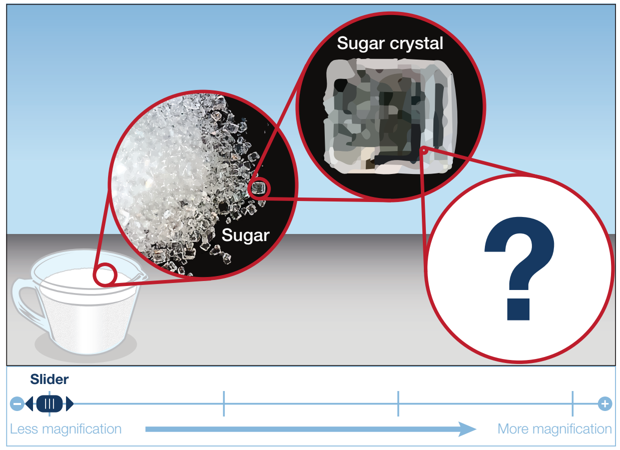
For an example of the slider functionality, visit http://www.numbersleuth.org/universe/magnify/.

Media (interactive image): The previous animation ends by zooming in to the sugar present in the measuring cup.

As the student slides the slider to the right, the image magnifies in a smooth transition. The student may also zoom out by sliding the slider to the left. The slider will automatically snap to the magnification points shown here, with the "?" being the highest magnification possible in this instance. Please note that the zoom circles shown here are for illustrative purposes only; the image for each magnification level will fill the entire student view in the platform.

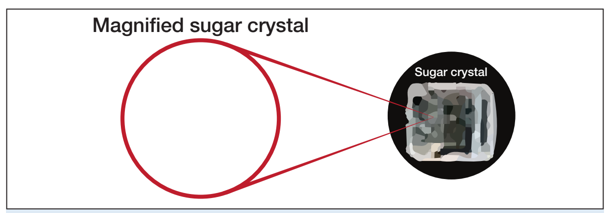
Slider

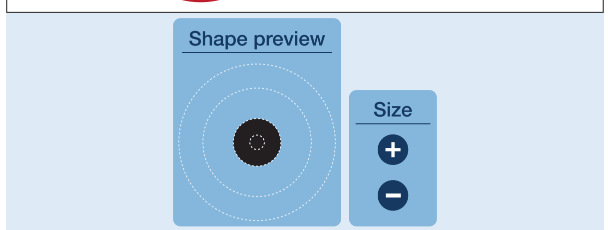
Platform prompt and student control



Part (a) Describe **one** way that the sugar and the water would look similar if they were both magnified under a powerful microscope.

Part (b) Develop a model to show what the sugar would look like when you slide the slider above to the greatest magnification. Change the size of the circle in the "Shape preview" box below to represent the matter that makes up the sugar, and drag one or more of the circles into the area for the magnified sugar crystal.





Click NEXT to continue to the next question. NEXT

5-PS1-1
Develop a model to describe that matter is made of particles too small to be seen.

Full alignment to the PE and targeted dimensions is intended through the entirety of the item cluster. Partial to strong alignment to the dimensions for each item is achieved through alignment to the evidence statements, and is inclusive of all item parts for any given item.

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.

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.

- Use models to describe phenomena.

Scale, Proportion, and Quantity

- Natural objects exist from the very small to the immensely large.

Part (a) Describe **one** way that the sugar and the water would look similar if they were both magnified under a powerful microscope.

Click NEXT to continue to the next question.

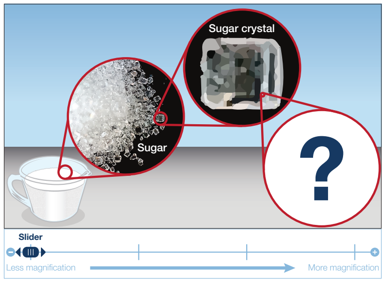
NEXT

Stem

Student types responses in this field. Field expands as needed.

Platform prompt and student control

Stimulus viewable by the student



Key

The sugar and water would look similar because you would see that they are both made up of many tiny particles.

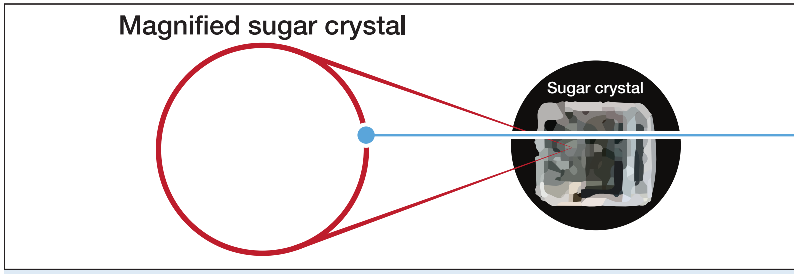
Item Type: Short Answer
Estimated Time: 2 min

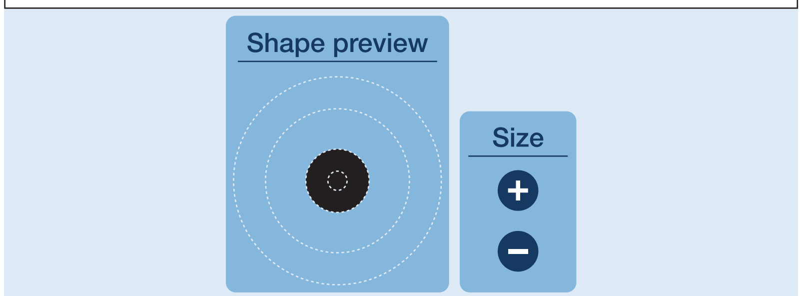
Evidence Statement Alignment:
(5-PS1-1)
(1) Components: (a) Students develop a model to describe a phenomenon, that includes the idea that matter is made of particles too small to be seen. In the model, students identify and describe the relevant components for the phenomenon, including: (ii) Particles of matter that are too small to be seen.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student demonstrates understanding that the sugar and water (or all matter) are made up of particles too small to be seen.

Scoring Notes: 1 point is awarded for a correct response; see Key.

Part (b) Develop a model to show what the sugar would look like when you slide the slider above to the greatest magnification. Change the size of the circle in the "Shape preview" box below to represent the matter that makes up the sugar, and drag one or more of the circles into the area for the magnified sugar crystal.

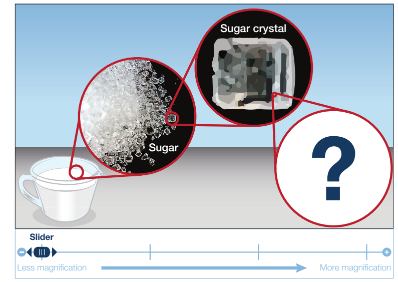




Click NEXT to continue to the next question.

NEXT

Stimulus viewable by the student



Stem

Drag-and-drop target

Object palette: The student can manipulate the size of the drag-and-drop object by clicking the "+" or "–" in this palette. Changes made to the object are reflected in the "Shape preview" area in real time (dashed lines are indicative of the possible sizes allowed by this palette, but will not be visible to the student). When ready, the student will drag and drop the object into the "Magnified sugar crystal" drag-and-drop target. The drag-and-drop object replenishes.

Platform prompt and student control

Item Type: Building a Model (Drag-and-Drop)
Estimated Time: 3 min

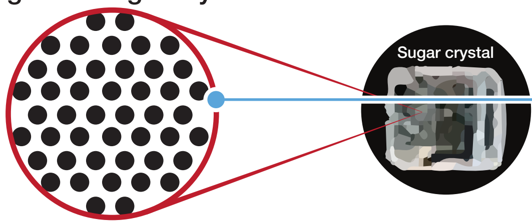
Evidence Statement Alignment:
(5-PS1-1)
(1) Components: (a) Students develop a model to describe a phenomenon that includes the idea that matter is made of particles too small to be seen. In the model, students identify and describe the relevant components for the phenomenon, including: (i) Bulk matter (macroscopic observable matter; e.g., as sugar, air, water).

(1) Components: (a) Students develop a model to describe a phenomenon, that includes the idea that matter is made of particles too small to be seen. In the model, students identify and describe the relevant components for the phenomenon, including: (ii) Particles of matter that are too small to be seen.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student can develop a model to describe how matter—sugar—is made up of many particles too small to be seen.

Part (b) Develop a model to show what the sugar would look like when you slide the slider above to the greatest magnification. Change the size of the circle in the “Shape preview” box below to represent the matter that makes up the sugar, and drag one or more of the circles into the area for the magnified sugar crystal.

Magnified sugar crystal



Shape preview

Size

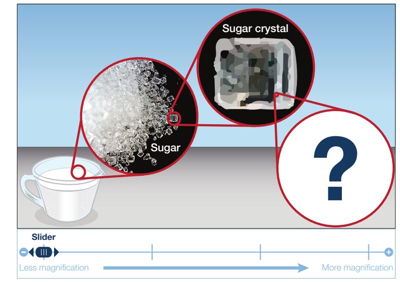
+

–

Click NEXT to continue to the next question.

NEXT

Stimulus viewable by the student



Stem

Drag-and-drop objects in target.

Object palette: The student can manipulate the size of the drag-and-drop object by clicking the “+” or “–” in this palette. Changes made to the object are reflected in the “Shape preview” area in real time. The drag-and-drop object replenishes.

Platform prompt and student control

Item Type: Building a Model (Drag-and-Drop)
Estimated Time: 3 min

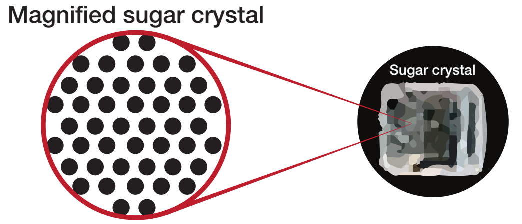
Evidence Statement Alignment:
(5-PS1-1)
(1) Components: (a) Students develop a model to describe a phenomenon that includes the idea that matter is made of particles too small to be seen. In the model, students identify and describe the relevant components for the phenomenon, including: (i) Bulk matter (macroscopic observable matter; e.g., as sugar, air, water).

(1) Components: (a) Students develop a model to describe a phenomenon, that includes the idea that matter is made of particles too small to be seen. In the model, students identify and describe the relevant components for the phenomenon, including: (ii) Particles of matter that are too small to be seen.

Note on Item Alignment:
What is being elicited from the student (evidence)? The student can develop a model to describe how matter—sugar—is made up of many particles too small to be seen.

Scoring Notes: Automated scoring is feasible, but the option to score a student’s response in conjunction with Item 5 is also viable. 1 point is awarded for a model that includes multiple (three or more) circles of the smallest or second to smallest size. The student does not need to include the shapes in a lattice-like arrangement to receive full credit. Students may use a number of different decision rules to build their models. Scoring of Item 4 Part (b) and Item 5 should be linked and should reflect this expectation.

YOUR MODEL:



Part (a) Your model shows what the sugar would look like if you could magnify it using a powerful microscope. Describe what your model shows about the sugar crystals that your eyes are unable to see.

Part (b) Describe how your model helps explain why the sugar seemed to disappear after the lemonade mixture was stirred.

Click NEXT to continue to the next question.

NEXT

5-PS1-1
Develop a model to describe that matter is made of particles too small to be seen.

Full alignment to the PE and targeted dimensions is intended through the entirety of the item cluster. Partial to strong alignment to the dimensions for each item is achieved through alignment to the evidence statements, and is inclusive of all item parts for any given item.

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.

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.

- Use models to describe phenomena.

Scale, Proportion, and Quantity

- Natural objects exist from the very small to the immensely large.

Part (a) Your model shows what the sugar would look like if you could magnify it using a powerful microscope. Describe what your model shows about the sugar crystals that your eyes are unable to see.

Stem

Note: Student's model created in Item 4 Part (b) is visible.

Student types responses in these fields. Fields expand as needed.

Part (b) Describe how your model helps explain why the sugar seemed to disappear after the lemonade mixture was stirred.

Key

Part (a) *The sugar crystals are made up of many small particles that you cannot see just by looking at them.*

Part (b) *The bigger sugar pieces you can see are made up of many particles that are too small to be seen, and the large sugar pieces separated into small particles when added to the liquid, so this is why the sugar could not be seen and looked like it disappeared after it was added to the lemonade mixture.*

Click NEXT to continue to the next question.

NEXT

Platform prompt and student control

Item Type: Constructed Response
Estimated Time: 6 min *(for Parts a and b)*

Evidence Statement Alignment:

(5-PS1-1)

(2) Relationships: (a) In the model, students identify and describe relevant relationships between components, including the relationships between: (i) Bulk matter and tiny particles that cannot be seen (e.g., tiny particles of matter that cannot be seen make up bulk matter). [Part (a)].

(3) Connections: (a) Students use the model to describe how matter composed of tiny particles too small to be seen can account for observable phenomena (e.g., air inflating a basketball, ice melting into water). [Part (b)].

Note on Item Alignment:

What is being elicited from the student (evidence)? The student can describe the relevant parts of the model to support the explanation that the sugar and the lemonade mixture are composed of particles too small to be seen.

Scoring Notes: 1 point is awarded for each correct response for Part (a) and for Part (b); see Key for samples of correct responses.

Item	Item Part	Brief Description	Item Type	PE	DCI	SEP	CCC	EV Level	EVs	Points	Estimated Time (min)	Hand or Automated Scoring
Stimulus		Preparing lemonade	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	3	N/A
1	1	Designing and populating a data table	Text Entry/ Table Fill-In	5-PS1-2	N/A	5	3	1	1.a.i 1.a.ii	2	2	A
2	2a	Calculate mass of ingredient	Computation	5-PS1-2	PS1.A PS1.B	5	3	1	1.a.i 1.a.ii	1	1	A
	2b	Graphing masses of ingredients	Graphing					2	2.a	2	2	A
	2c	Describe properties of individual ingredients	Short Answer					2	2.c	1	2	H
3	3a	Claim for conservation of mass	Multiple Choice	5-PS1-2	PS1.A PS1.B	5	3	2	2.d	1	1	A
	3b	Identify evidence of conservation of mass	Multiple Select					2	2.d	1	1	A
Stimulus		Investigating ingredients	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	N/A
4	4a	Describe that both sugar and water are made up of particles	Short Answer	5-PS1-1	PS1.A	2	3	1	1.a.ii	1	2	H
	4b	Building a model to show particles of matter	Building a Model (Drag-and-Drop)					1	1.a.i 1.a.ii	1	3	A or H
5	5a–b	Describing the model and use of model in explaining science phenomenon	Constructed Response	5-PS1-1	PS1.A	2	3	2, 3	2.a.i 3.a	2	6	H
								Total:	**9 of 11**	**12**	**24**	

Evidence Statements for: 5-PS1-1

✓ Included (Item 4b)
1. Components of the model (a) Students develop a model to describe a phenomenon that includes the idea that matter is made of particles too small to be seen. In the model, students identify the relevant components for the phenomenon, including: i. Bulk matter (macroscopic observable matter; e.g., as sugar, air, water).

✓ Included (Item 4a–b)
1. Components of the model (a) Students develop a model to describe a phenomenon that includes the idea that matter is made of particles too small to be seen. In the model, students identify the relevant components for the phenomenon, including: ii. Particles of matter that are too small to be seen.

✓ Included (Item 5a)
2. Relationships (a) In the model, students identify and describe relevant relationships between components, including the relationships between: i. Bulk matter and tiny particles that cannot be seen (e.g., tiny particles of matter that cannot be seen make up bulk matter).

Not Included
2. Relationships (a) In the model, students identify and describe relevant relationships between components, including the relationships between: ii. The behavior of a collection of many tiny particles of matter and observable phenomena involving bulk matter (e.g., an expanding balloon, evaporating liquids, substances that dissolve in a solvent, effects of wind).

✓ Included (Item 5b)
3. Connections (a) Students use the model to describe how matter composed of tiny particles too small to be seen can account for observable phenomena (e.g., air inflating a basketball, ice melting into water).

Evidence Statements for: 5-PS1-2

✓ Included (Items 1, 2a)
1. Representation (a) Students measure and graph the given quantities using standard units, including: i. The weight of substances before they are heated, cooled, or mixed.

✓ Included (Items 1, 2a)
1. Representation (a) Students measure and graph the given quantities using standard units, including: ii. The weight of substances, including any new substances produced by a reaction, after they are heated, cooled, or mixed.

✓ Included (Item 2b)
2. Mathematical/computational analysis (a) Students measure and/or calculate the difference between the total weight of the substances (using standard units) before and after they are heated, cooled, and/or mixed.

Not Included
*2. Mathematical/computational analysis (b) Students describe the changes in properties they observe during and/or after heating, cooling, or mixing substances.

✓ Included (Item 2c)
2. Mathematical/computational analysis (c) Students use their measurements and calculations to describe that the total weights of the substances did not change, regardless of the reaction or changes in properties that were observed.

✓ Included (Item 3a–b)
2. Mathematical/computational analysis (d) Students use measurements and descriptions of weight, as well as the assumption of consistent patterns in natural systems, to describe evidence to address scientific questions about the conservation of the amount of matter, including the idea that the total weight of matter is conserved after heating, cooling, or mixing substances.

Excerpts from the SAIC Assessment Framework

The following sections of this document were taken directly from the SAIC Assessment Framework and are included to provide a high-level overview of item cluster expectations. Assessment Framework page references are provided in parentheses. The SAIC Assessment Framework and Item Specifications Guidelines, respectively, can be accessed, in their entirety, at the following locations:

- http://www.ccsso.org/Resources/Publications/Science_Assessment_Item_Collaborative_Assessment_Framework.html
- http://www.csai-online.org/spotlight/science-assessment-item-collaborative

Preface (p. 1)

The Science Assessment Item Collaborative (SAIC) Assessment Framework ("Assessment Framework") provides a range of options and accompanying rationales for the development of Next Generation Science Standards (NGSS)-aligned items and summative assessments. The Assessment Framework is designed to be used in concert with the *Item Specifications Guidelines* to aid state education agencies (SEAs) and other entities in documenting the processes needed to drive the development of NGSS-aligned items and assessments. Due to the interrelated nature of the documents, elements of the Assessment Framework that specifically detail the characteristics of the assessments and associated development considerations may also appear in the *Item Specifications Guidelines*.

The Assessment Framework principally draws on the following three seminal resources:

- the National Research Council (NRC)'s *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (NRC, 2012), hereafter referred to as the "*K–12 Framework*";
- the Next Generation Science Standards: For States, by States (NGSS Lead States, 2013), hereafter referred to as the "NGSS"; and
- the NRC Board on Testing and Assessment (BOTA)'s report *Developing Assessments for the Next Generation Science Standards* (NRC, 2014), hereafter referred to as the "BOTA report."

The research-supported recommendations and evidence base for practice that are embodied in these reports are foundational to the approach to development of next-generation science assessments (NGSAs) that is endorsed in the Assessment Framework.

Chapter One: Introduction (p. 4)

A new approach to K–12 science education was presented in the National Research Council (NRC)'s *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (NRC, 2012). The *K–12 Framework* articulates a broad set of rigorous expectations to support all students in achieving scientific literacy, and provided guidelines on how to prepare students to be able to pursue science, technology, engineering, and mathematics (STEM) careers. The *K–12 Framework* organizes science learning around three main dimensions: Science and Engineering Practices (SEPs), Crosscutting Concepts (CCCs), and Disciplinary Core Ideas (DCIs). It emphasizes that these three dimensions must be interwoven into every aspect of science education—curriculum, instruction, and assessment—in order to fully achieve the vision set forth for science education.

The *K–12 Framework* was subsequently used as the research-supported foundation for the development of the NGSS. The NGSS and the *K–12 Framework* present a holistic approach to science education in maintaining that students must, in both instruction and assessment, "engage in scientific and engineering practices in the context of disciplinary core ideas, and make connections across topics through the crosscutting ideas" (NRC, 2014, p. 4). This multidimensional approach to science education presents both opportunities and challenges for states as they begin to implement the rigorous new standards and develop NGSS-aligned assessments.

Chapter One: An Assessment Framework (pp. 4–6)

A primary purpose of the SAIC is to support states in the development of a pool of high-quality items for large-scale summative assessment. To achieve this goal, the SAIC is initially developing guidance documents outlining a systematic, methodical, and research-based approach to the design and development of NGSS-aligned summative assessments. This approach begins with the development of an assessment framework, aimed at state science assessment coordinators and assessment developers, and serving as a bridge between the NGSS and methods of assessing those standards.

[. . .] It is anticipated that states will use the Assessment Framework for a number of purposes. Along with the Item Specifications Guidelines, the Assessment Framework will be the guiding document to inform the development of requests for proposals (RFPs) that will be used to select and guide assessment vendors in the development of NGSS-aligned assessments. It will also serve as a guiding document for the development of state and local test specifications and blueprints. The Assessment Framework may also be a valuable communication tool, providing information to key stakeholders and professional development providers.

Chapter One: Assessment Framework Development Process (pp. 7–8)

Development of the NGSS-based Assessment Framework described in this report was led by the SAIC, with WestEd as the primary author. This effort entailed a comprehensive state survey, multiple rounds of member and expert reviews, and strategic refinement of the emerging recommendations. Members of Achieve Inc.—including Dr. Stephen Pruitt, Senior Vice President of Achieve Inc., who coordinated development of the NGSS—and other experts in assessment design and psychometrics provided valuable feedback on drafts of the Assessment Framework and provided consultation during its development. Further details on the development process can be found in Appendix G.

The members of the SAIC are a diverse group of states and other jurisdictions. Some are members of one of the two major Race to the Top assessment consortia (the Smarter Balanced Assessment Consortium [Smarter Balanced] and the Partnership for Assessment of Readiness for College and Careers [PARCC]); others do not participate in either consortium. Some members adopted the NGSS by name, while others adopted the full NGSS but with a name change or opted for a partial adoption. Some members plan for a full computer-based administration of NGSS-based assessments, while others plan to use a mix of computer-based delivery and paper-and-pencil delivery. Finally, some members fully embrace the recommendations in the BOTA report (NRC, 2014), while other members support a more state-mediated approach to the transition to the NGSS. Despite these differences, all members of the SAIC have worked together to achieve the goals that they established as a team.

Chapter One: Context of the Assessment Framework (pp. 9–10)

The primary focus of this Assessment Framework is to build a basis of item development for NGSS large-scale assessment within the context of overall test design. The Assessment Framework should be considered a starting point for the implementation of a large-scale assessment measuring the NGSS, rather than being considered the final model. It should be noted that the item cluster model presented in the Assessment Framework has not been developed and fully implemented in a state testing system for science, although significant parts of it have. Lessons learned through large-scale development will present opportunities to adjust the model presented and tools recommended. The descriptions and expectations presented in the Assessment Framework should be considered a starting point, rather than the definitive end product. In addition, there are psychometric challenges that will need to be addressed (and limits pushed) for tests built using the item cluster model as the basic building component. These issues include acceptable content coverage, pilot testing, score generalizability, and number of score points to achieve reporting expectations. Matrix sampling is considered an important test design consideration for achieving a reasonable amount of content coverage and for achieving aggregate level reporting at the school, district, and state levels. In addition, reporting for the individual student for anything other than overall science ability will be problematic to support using only item clusters. Even for overall science ability at the individual student level, individual reliability of scores may not be as strong as is achievable with a test composed primarily of individual items. The acceptable limits for the described concerns will need to be addressed and determined by individual states through their development and implementation efforts.

The Assessment Framework's focus on large-scale assessment was an outcome of needs expressed by states to begin the conversation about how to develop such an assessment while still being true to the principles and expectations of the *K–12 Framework* (NRC, 2012), the NGSS (NGSS Lead States, 2013), and the BOTA report (NRC, 2014). The presentation of this Assessment Framework in no way espouses the use of a single test in isolation to measure and report on the full NGSS. The BOTA report (NRC, 2014) describes a comprehensive assessment program approach. Concerns previously described in this section can be lessened if the assessment is used within the context of an assessment system that provides (through other assessments) information with greater usability and generalizability at the school level.

This Assessment Framework is not intended to provide a full assessment solution for states. Its intent is to present an acceptable solution for achieving alignment to the NGSS for large-scale assessment. Many lessons remain to be learned as this solution is pursued.

Chapter Four: Context of Item Clustering (pp. 21–22)

SAIC members have come to an agreement on common terminology used to describe two components of this emerging assessment. First, an *item cluster* is a set of items (usually between four and six items, with some items having more than one part) that are based on at least one common stimulus (e.g., text, audio, video, animation, simulation, experiment). Individual items that are part of an item cluster are not intended to be separated and used independently from one another. Second, because the term *performance-based task* can be used to describe a broad family of assessment activities, the SAIC has adopted the definition outlined by Smarter Balanced: "[a] performance task involves significant interaction of students with stimulus materials and/or engagement in a problem solution, ultimately leading to an exhibition of the students' application of knowledge and skills" (Smarter Balanced, 2012, p. 1).

The BOTA report (NRC, 2014) recommends the use of assessment tasks with multiple components, rather than more traditional, discrete, stand-alone items:

> *CONCLUSION 2-1 Measuring the three-dimensional science learning called for in the framework and the Next Generation Science Standards requires assessment tasks that examine students' performance of scientific and engineering practices in the context of crosscutting concepts and disciplinary core ideas. To adequately cover the three dimensions, assessment tasks will generally need to contain*

multiple components (e.g., a set of interrelated questions). It may be useful to focus on individual practices, core ideas, or crosscutting concepts in the various components of an assessment task, but, together, the components need to support inferences about students' three-dimensional science learning as described in a given performance expectation. (p. 44)

As presented in the Assessment Framework, item clusters are the large-scale summative assessment fulfillment of the assessment tasks recommended in the BOTA report. Item clustering will be needed in order to fully and accurately assess the NGSS. Additionally, each item within an item cluster must be aligned to at least two dimensions of the NGSS, with a strong preference that every effort be made, when feasible, to develop items aligned to all three dimensions of the NGSS. The overall item cluster must demonstrate alignment to all three dimensions.

One concern with an item-cluster-only approach for item development is that if the item cluster is appropriately developed, extracting individual items for stand-alone use will not be possible due to the scaffolded and intertwined nature of the items. A final consideration is that, at present, there are no known extant NGSS items developed that are fully aligned to the NGSS. For this reason, one planned outcome of the SAIC work is two prototype item clusters. These prototypes, which will be made available as separate documents, will offer examples of the item-cluster and NGSS alignment expectations.

The Assessment Framework presents an approach to item development that takes into consideration the following premises:

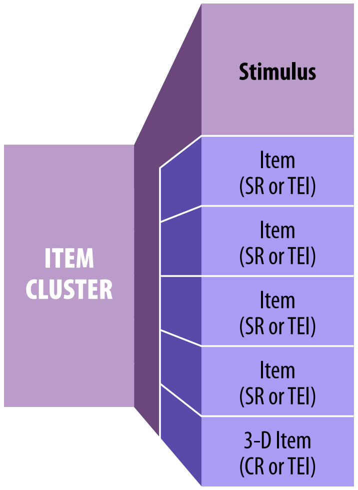
- Item clusters, not individual items, are the base unit for the SAIC test development. That is, individual items are intentionally developed to be situated within the context of an item cluster and not to be used as stand-alone items.
- Item clusters are the primary focus for developers in terms of alignment to the NGSS. That is, each item cluster must demonstrate strong three-dimensional alignment to the NGSS.
- To qualify as NGSS-aligned, item clusters must be aligned to one or more PEs and must be inclusive of all of the dimensions associated with the PE(s) (i.e., DCI, SEP, CCC).
- Each individual item within the cluster must align with at least two dimensions of the NGSS (e.g., DCI, SEP, and/or CCC) to qualify for inclusion in an item cluster.

Item clusters as described in this chapter and in the Item Specifications Guidelines fulfill these expectations.

Chapter Four: Architecture of Item Clusters (pp. 22–25)

As shown in Figure 1, the basic structure of an item cluster includes a common stimulus with an associated set of items.

Figure 1. Sample representation of the relationship of an item cluster to its component items



Each item is inextricably linked to the stimulus and to the other items within the item cluster, and the stimulus may be interspersed among the items to add information as needed. This means that student exposure to the stimulus is considered essential in order to respond correctly to any individual item, and that the item cluster must be constructed in such a way that individual performance on each item is adversely affected if an item is responded to without the context of the other items in the cluster. Testing time for each item cluster will be content dependent, but an approximation of 20 minutes of testing time per item cluster should be assumed. This time limit will allow for a reasonable overall test length while still providing an acceptable coverage of NGSS standards (i.e., PEs).

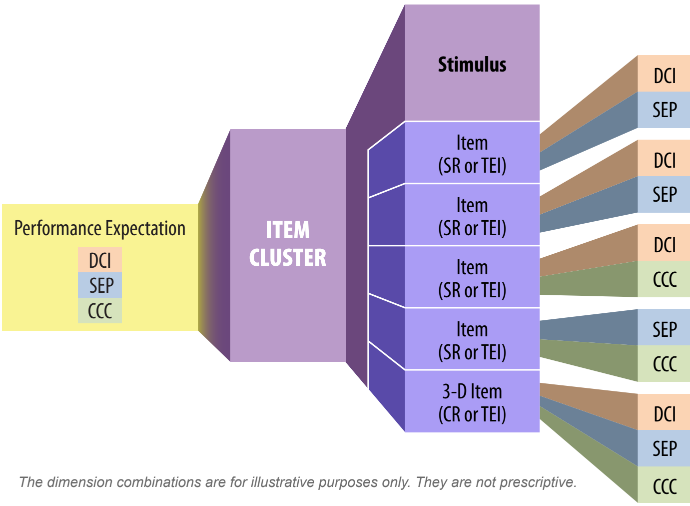
A stimulus is defined as a component of the cluster that does not directly require a student response. A stimulus can include one or more of the following:

- text;
- audio;
- video;
- discussion;
- animation/simulation;
- experimentation;
- activity; and/or
- demonstration.

Initially, stimuli will be identified or developed with the intent of inclusion on a large-scale summative assessment. States and developers should pursue creative solutions and should not allow current challenges of administration to constrain their thinking. The item cluster model is designed to allow for gradual evolution of stimuli, but still maintain NGSS alignment expectations.

An example of an item cluster's overall three-dimensional alignment is shown in Figure 2, with the dimensions of each item in a simplified single-PE cluster included.

Figure 2. Sample representation of the relationship of an item cluster aligned to a single PE to its component items, with item-aligned dimension combinations shown



The dimension combinations are for illustrative purposes only. They are not prescriptive.

- It should be noted that all items will exhibit some degree of alignment to the disciplinary context of the DCI, as all items are inextricably linked to the context, which was selected to align to the discipline(s) associated with the PEs. Therefore, every item in an item cluster will naturally fall within the content limits of the DCI, but not every item may truly call for the assessment of understanding of the content put forth in the DCI. Thus, items that only align to SEPs/CCCs are not intended to be viewed as devoid of a disciplinary context, but, rather, are intended to be viewed as items that place relatively greater emphasis on assessing an associated SEP and/or CCC than they do on assessing the underlying DCI content. Each SEP and CCC has its own knowledge that is most relevant in context of a DCI.
- If an evidence statement appears to align to a single SEP or CCC dimension, it is recommended that the evidence statement be grouped with the DCI, in order to prevent an item writer from developing an item to a single dimension in isolation (e.g., attempting to assess a science practice in isolation, without tying the item to the context and/or the DCI).
- At least one item should be aligned to all three dimensions, as shown in Figure 2, as this is the overall vision of the NGSS.
- Each item is inextricably linked to the stimulus and to the other items within the item cluster. This means that student exposure to the stimulus is considered essential in order for the student to respond correctly to any individual item, and that the cluster of items must be constructed in such a way that individual performance on each item is adversely affected if an item is responded to without the context of the other items in the cluster. (See the Item Specifications Guidelines for more information on stimuli for item clusters.)
- Testing time for each item cluster will be content dependent, but an estimate of 20 minutes of testing time per item cluster is assumed for summative assessment purposes. This estimate will be further refined as prototypes are completed.
- Each item cluster will have items tied to evidence statement selections for one or more PEs. These evidence statement selections are the fundamental component of item alignment with scientific content. Item clusters aligned to more than one PE could be from the same domain (i.e., Physical Sciences, Life Sciences, Earth and Space Sciences), but could also be from related, but different, content areas (e.g., photosynthesis and chemical reactions). PEs can also be from different domains. PEs from the domain of Engineering, Technology, and Applications of Science should always be bundled with PEs from one of the science disciplines.

  The rationale for correlating the parts of a PE evidence statement with two or more of the PE's dimensions is that such a correlation provides a building block for item construction when the PE is bundled with one or more other PEs in an item cluster. Looking at the entirety of the dimensions and evidence statements for two or more PEs in an item cluster can be somewhat overwhelming in terms of the amount of information provided in relation to assessment goals. By structuring the PE and evidence statement components into natural dimensional/evidence-statement relationships that might form the basis of an item in an item cluster, the item cluster developer can better perceive how all of these PE elements fit together and how they might be used, along with the multidimensional alignment groupings for other PEs in an item cluster, to form a balanced, conceptually cohesive item cluster.
- While it may be possible to develop items within a single cluster that are collectively sufficient to assess the entirety of a single PE, this is not preferable and will not be possible in many, if not most, cases. For item clusters inclusive of more than one PE, it is not expected that a single item cluster will be able to provide the opportunity for a student to generate evidence of every aspect of each PE in the item PE bundle.

More detailed explanations of item clusters are provided in the Item Specifications Guidelines.

References

National Research Council (NRC). (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Committee on Conceptual Framework for the New K–12 Science Education Standards. Board on Science Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (NRC). (2014). *Developing assessments for the Next Generation Science Standards*. Committee on Developing Assessments of Science Proficiency in K–12. Board on Testing and Assessment and Board on Science Education, J. W. Pellegrino, M. R. Wilson, J. A. Koenig, & A. S. Beatty (Eds.). Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Next Generation Science Standards (NGSS) Network. (2015a). NGSS Evidence Statements Executive Summary. Retrieved from http://www.nextgenscience.org/sites/ngss/files/Exec Summary_NGSS Evidence Statements Front Matter Jan 2015.pdf

Next Generation Science Standards (NGSS) Network. (2015b). *NGSS Grades K–5 Evidence Statements*. Retrieved from http://www.nextgenscience.org/k-5-evidence-statements

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: Achieve, Inc.

Smarter Balanced Assessment Consortium (Smarter Balanced). (2012). Smarter Balanced Assessment Consortium: Performance task specifications. Retrieved from http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/05/TaskItemSpecifications/PerformanceTasks/PerformanceTasksSpecifications.pdf