AIM User Manual
Properties of and Changes in Matter
Elementary School Student Assessment

Overview

The AIM Properties of and Changes in Matter Elementary School Student Assessment is a 25-item multiple-choice assessment developed for upper elementary grades science students. The assessment is based on the *Science Framework for the 2009 National Assessment of Educational Progress* (National Assessment Governing Board, 2008) and measures understandings of concepts in two related content areas.

1. Properties of matter
    - physical properties common to all objects and substances;
    - physical properties common to solids, liquids, and gases; and
    - chemical properties and particulate nature of matter; and the Periodic Table of the Elements.

2. Changes in matter
    - changes of state; and
    - physical and chemical changes and conservation of mass.

This user manual describes the background, development, measurement properties, and appropriate uses of the assessment. User manuals for other AIM assessments may be found at http://www.horizon-research.com/aim/instruments/.

Background

Horizon Research, Inc. (HRI) developed the AIM Properties of and Changes in Matter Student Assessment as part of a larger study. The project—Assessing the Impact of the MSPs: K–8 Science (AIM) was funded by the National Science Foundation under Grant no. DUE-0928177.[1] One goal of AIM was to develop instruments that researchers could use to study the theory of action that underlies much professional development for science teachers. Briefly, the model asserts that changes in teacher knowledge lead to changes in classroom practice (mediated by instructional materials), and ultimately, changes in student learning (see Figure 1). Despite the prominent role this model plays in professional development design, it has not been studied systematically, in part because of a lack of instruments. Among other products, AIM developed pairs of assessments—one for teachers and one for students—focused on the same science content area. These pairs of assessments enable the study of relationships between teacher

[1] Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

knowledge and student learning in specific science contexts. AIM assessments exist for four content areas: (1) evolution and diversity of life; (2) force and motion (Newton's first and second laws); (3) populations and ecosystems; and (4) properties of and changes in matter . For each content area, separate pairs of assessments were developed for upper elementary school and middle school levels.

Professional Development Theory of Action

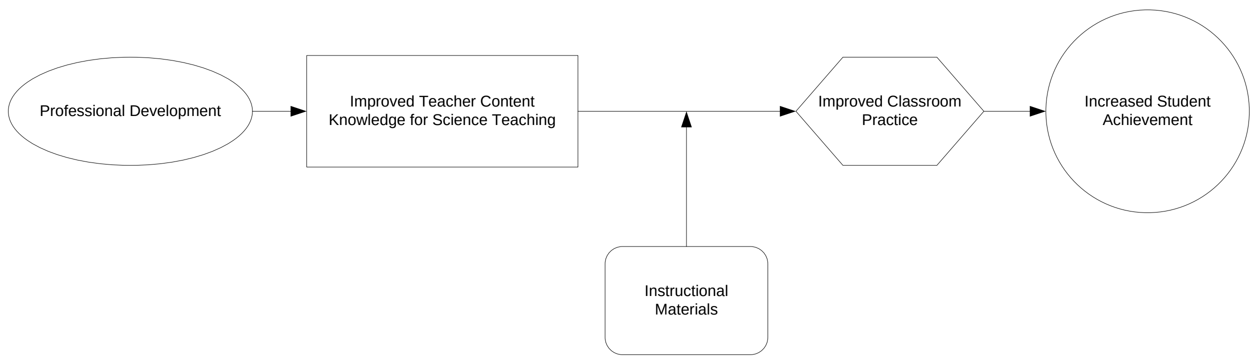

Figure 1

To enable large-scale research, HRI set out to create assessments that would be minimally burdensome, both for the test-taker and the researcher. Accordingly, HRI opted for a multiple-choice format, recognizing the limitations of such items. For instance, well-constructed, open-ended items may probe more depth of understanding than multiple-choice items, but they are more burdensome for both the researcher (in terms of scoring costs) and the test-taker (in terms of time required to complete the assessment). In addition, scoring open-ended items requires the training of raters to establish inter-rater reliability.

Development of the Properties of and Changes in Matter Elementary School Student Assessment

As described above, this development effort was part of a much larger and well-funded project, which afforded a thorough development process (see Figure 2).

Assessment Development Process

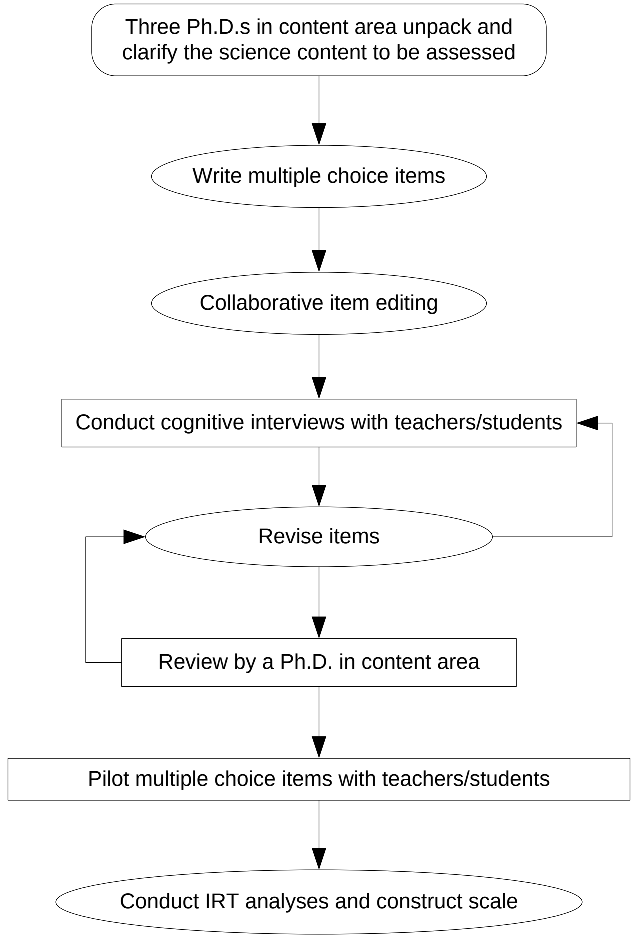

Figure 2

Clarifying the Content Domain

Development began with identifying the target content for the properties of and changes in matter assessments. We used the 2009 *NAEP Framework* for direction on the content of the AIM assessments. The *NAEP Framework* was based primarily on the *National Science Education Standards* (National Research Council, 1996) and the *Benchmarks for Science Literacy* (American Association for the Advancement of Science, 1993), but also reflected developments in science and policy that have taken place since those documents were published. HRI specified the assessment domain using two strands in the *NAEP Framework*: (1) properties of matter and (2) changes in matter. This process had three chemists/chemistry educators "unpack" the content into series of "sub-ideas" for upper elementary students. These are the ideas that were considered in developing the elementary school student assessment. The final description of the content domain is shown in Table 1.

Table 1a
Properties of and Changes in Matter Content Domain

Properties of matter. Physical properties common to all objects and substances and physical properties common to solids, liquids, and gases; chemical properties, particulate nature of matter, and the Periodic Table of the Elements

Sub-ideas:

- Materials have properties.
  - The properties of anything are the characteristics used to describe that thing, such as color, weight (mass), size, and so on.
  - Various objects and substances (materials) can be identified and distinguished by their properties.
  - Weight (mass) and volume are properties that can be measured using appropriate tools. (For elementary school assessments, the more familiar "weight" is used rather than "mass.")
  - Properties of a sample may be measured (quantitative properties) or described using the five senses (qualitative properties).
  - Weight is the property of matter that makes it resist efforts to move or lift it and is usually measured with scales or balance. Qualitatively, the weight of a quantity of matter is often characterized as "heavy" or "light."
  - Volume is the property of matter that tells us how much space the matter occupies. Qualitatively, the volume of a quantity of matter is often characterized as "large" or "small."
- Objects (materials) vary in the extent to which they absorb and reflect light and conduct heat (thermal energy) and electricity.
  - Different objects absorb different colors of light that strikes them and reflect other colors, which give the objects the colors we observe.
  - Some objects allow electricity to easily pass through them (conductors) and others do not (insulators).
  - Some objects allow heat (thermal energy) to easily pass through them and others do not.
- Permanent magnets can repel or attract other permanent magnets.
  - Permanent magnets are bipolar. One pole is usually called the "north pole" and the other the "south pole." The unlike poles of permanent magnets attract one another and the like poles repel.
- Permanent magnets can also attract certain materials (induced magnets) at a distance.
  - Some materials (induced magnets) are attracted to magnets.
  - Some materials (non-magnets) are not attracted to magnets.
- Materials exist in several different states; the most common states are solid, liquid, and gas. Each state of matter has characteristic properties.
- Samples of materials can be classified by their behavior into solids, liquids, and gases.
  - Solids have a definite shape and volume that cannot be easily changed. (For elementary school assessments "certain shape" or "certain size" are used rather than "definite shape.")
  - Liquids flow to take the shape of their container and have a definite volume that cannot be easily changed.
  - Gases expand to fill any shape and volume container and can easily be compressed into a smaller volume container.
- Some objects are composed of a single material; others are composed of more than one material.
  - Some materials can be separated (decomposed) into two or more different materials.
  - Some materials cannot be separated (decomposed) into different materials.

Table 1b
Properties of and Changes in Matter Content Domain

Changes in Matter: Changes of state; physical and chemical changes and conservation of mass

Sub-ideas:

- One way to change matter from one state to another and back again is by heating and cooling.
  - Melting and freezing are terms for opposite changes in state, melting is solid → liquid; freezing is liquid → solid).
  - The melting point is the temperature at which a solid melts to form the liquid state of the substance. The freezing point of a liquid is the temperature at which the liquid freezes to form the sold state of the substance. For any particular substance (material), melting and freezing points are the same.
  - Heating a solid can raise its temperature to the melting point and continued heating will melt the solid.
  - Cooling a liquid lowers its temperature to the freezing point and continued cooling will freeze the solid.
  - Boiling (vaporizing) and condensing are opposite state-change processes. (Although similar statements can be made about the liquid-to-gas and gas-to-liquid changes of state, the concepts of vapor pressure and atmospheric pressure are involved and these are not suitable at this level, so it is just as well to say that boiling and condensing are similar to melting and freezing and not try to define boiling quantitatively.

Item Development

HRI staff drafted items individually then met to edit them collaboratively. As the pool of items grew, we began recruiting upper elementary school students for telephone cognitive interviews. We typically interviewed three students on each item in the pool using the interview protocol shown in Figure 3. After a round of interviews, HRI staff met to discuss students' responses and feedback. If substantive edits were made to an item, we interviewed additional students about the revised version. When interviews suggested no further edits were needed, we asked a content expert to review all of the items in the pool for content accuracy.

AIM Student Assessment Items
Cognitive Interview Protocol

Prologue Script:

We are developing test questions for elementary school students who have been studying properties of and changes in matter, and we need your help to get the questions just right. I realize that you may not have studied some of this yet in school, and I don't expect you to get all of the answers right. If you get a few wrong, it will help me know whether we have written the answer choices well. You can ask me to explain any words or situations that may be unfamiliar or confusing, but I can't give you the answer to any of the questions until the end of the interview. Remember, the point is to help us write a good test, not to test what you do or don't know. You won't get a grade or anything like that on the test. Do you have any questions before we get started? If at any point in the interview you would like to stop, just say so.

Procedure:

- Ask student to read aloud and "think aloud" as they read the questions and answer choices, if they are comfortable doing so. Remind the student to go back and reread the question to himself/herself if he/she needs to. If reading the question aloud is too distracting or uncomfortable, allow the student to read the question to himself/herself.

- It is not necessary to time how long it takes for the student to arrive at an answer, but if it takes an especially long time on a question, please make a note of it in the comment area of the notes.

- For each item, ask:
  1. Why did you choose that answer? (probe for words or diagrams they keyed in on, as well as their thinking behind the response)
  2. What did you think of each of the other answer choices?
  3. Was there an answer choice you were expecting to see, but did not? What was it?
  4. Were there any words or diagrams you did not really understand, or situations that made the question confusing?
  5. Is there anything about the question that did not confuse you, but that you think might confuse other students?
  6. Do you have any other comments on the item?

Figure 3

An example student assessment item resulting from this process is shown in Figure 4. (correct answer is B)

Properties of and Changes in Matter Item

A student does an experiment to decide if two objects are magnets. She holds them near a magnet and notices Object 1 sticks to the magnet, but the magnet pushes Object 2 away.

Based on the experiment, what does the student know about the two objects?

A. Object 1 is a magnet, but Object 2 may or may not be a magnet.
B. Object 1 may or may not be a magnet, but Object 2 is a magnet.
C. Object 1 and Object 2 are magnets.
D. Object 1 and Object 2 may or may not be magnets.

Figure 4

This item illustrates some features common to all AIM student assessment items. It is not included in the AIM assessment, but is shown here to illustrate item features. This example item may be flawed and is not intended to be used in any assessments. As mentioned previously, all items are multiple choice. All include only four choices and preclude as choices "none of the above," "all of the above," or multiple correct answers such as, "A and B but not C."

Pilot

We selected 35 items to pilot with approximately 702 students of teachers recruited from mailing lists of elementary grades teachers across the country. The pilot was administered as a paper form by recruited teachers.

Table 2
Characteristics of the Pilot Test Sample

	Percent of Students
Grade Level	
4th grade	28
5th grade	72
English is primary language	
Yes	88
No	12
Gender	
Female	46
Male	54
Race/Ethnicity†	
American Indian or Alaskan Native	3
Asian	2
Black or African American	15
Hispanic or Latino	23
Native Hawaiian or Other Pacific Islander	1
White	67

† Percentages may add up to more than 100 as students could select multiple categories.

Measurement Properties of the Assessment

Following is a description of the content coverage of the assessment, information about the validity and reliability of the assessment, and the results of the item response theory (IRT) analysis.

Content Coverage

Using results from the pilot, 25 items were selected for the final form. The distribution of items by sub-idea is shown in Table 3. There are fewer sub-ideas in Table 3 than in the content unpacking (see Table 1), as limiting the assessment to a total of 25 items required restricting the coverage of sub-ideas. In some cases a sub-idea may not be represented in the final assessment because it was deemed to be less central than others. In other cases, items associated with the sub-idea did not perform as well as others in the pilot study.

Table 3
Number of Items Addressing Each Sub-Idea

Sub-Ideas:	Number of Items
A. Materials have properties.	3
B. Weight (mass) and volume are properties that can be measured using appropriate tools.	3
C. Objects (materials) vary in the extent to which they absorb and reflect light and conduct heat (thermal energy) and electricity.	2
D. Materials exist in several different states; the most common states are solid, liquid, and gas. Each state of matter has characteristic properties.	2
E. Samples of materials can be classified by their behavior into solids, liquids, and gases.	5
F. Some objects are composed of a single material; others are composed of more than one material.	1
G. Materials can be either permanent magnets, induced magnets, or non-magnets. Permanent magnets can repel or attract other permanent magnets.	2
H. Permanent magnets can also attract certain materials (induced magnets) at a distance.	2
I. One way to change matter from one state to another and back again is by heating and cooling.	6
J. Matter is "stuff", everything we can see and/or touch.	2

Table 4 shows the answer key and content association for each item on the assessment. The letter "P" denotes a primary association with the sub-idea being targeted by the item. An "S" denotes a secondary association with a sub-idea that is also necessary in order to answer the item correctly, but is not the primary idea being assessed.

Table 4
Answer Key and Sub-Idea Associations

Item #	Key	Sub-Idea A	B	C	D	E	F	G	H	I	J
1	D				P						
2	B									P	
3	B					P					
4	D									P	
5	A	P									
6	D										P
7	D	P									
8	B				P					S	
9	B							P	P		
10	C										P
11	C			P							
12	B					P					
13	B			P							
14	A	P									
15	C					P					
16	D					P					
17	C						P				
18	D									P	
19	A		P								
20	D					P					
21	C		P								
22	C							P	P		
23	B									P	
24	A		P								
25	A									P	
Primary:		3	3	2	2	5	1	2	2	5	2
Secondary:		0	0	0	0	0	0	0	0	1	0
Total:		3	3	2	2	5	1	2	2	6	2

Validity
Three lines of evidence support the argument that the assessment is a valid measure of students' knowledge of these properties of and changes in matter ideas. First, cognitive interviews with students established that students interpret the items as intended and that they must use their knowledge of content to answer the items correctly. Second, a content expert (individual with a Ph.D. in chemistry) reviewed the assessment items to ensure content accuracy. Third, factor analysis indicates that all items on the assessment measure a single dominant trait. HRI termed this trait "content knowledge about properties of and changes in matter."

Reliability
Both classical test and item response theory (IRT) analyses were conducted on the pilot data and those results were used to select items for the final assessment. The assessment has an IRT reliability of 0.77; reliabilities above 0.60 are generally considered acceptable for making judgments about groups (higher reliabilities are required for making high-stakes decisions about individuals.

Speededness
In the pilot, teachers were instructed to give their students 50 minutes or the length of the class period (whichever was shorter) to complete the test. There was no evidence of speededness.

Using the Assessment

The AIM Properties of and Changes in Matter Student Assessment is available at no cost through an online process to those who agree to the terms of use (see the Appendix). To complete the terms of use agreement, visit http://www.horizon-research.com/aim/instruments/.

Appropriate Use
The AIM Properties of and Changes in Matter Student Assessment yields a score for each individual. However, the assessment is not valid for making judgments about individuals based on those scores. For instance, assigning student grades based on scores is not a valid use of the assessment. The assessment was not validated for such purposes.

HRI developed the assessment for use in research contexts involving groups of teachers. Appropriate uses with sufficiently large groups of teachers (20 or more) include:

- Measuring the change in group mean from pre-workshop to post-workshop;
- Comparing the gains of treatment and control groups; and
- Researching the relationship between teacher knowledge and other variables (e.g., student learning).

Amount of Time Required to Complete the Assessment
Although there is no evidence of speededness, it is recommended that at least 45 minutes be allowed for completing the assessment.

Computing Scores

Scores may be computed either as number correct or percent correct. Results of an item-response theory (IRT) analysis are shown in Table 5. This table can be used to convert a raw score in terms of number correct to the corresponding scale score.

Table 5
Assessment Score Conversions

Raw Score	Scale Score
0	0
1	13
2	21
3	26
4	30
5	33
6	36
7	39
8	41
9	43
10	45
11	47
12	49
13[†]	51[†]
14	53
15	55
16	57
17	59
18	61
19	64
20	66
21	70
22	73
23	79
24	87
25	100

[†] Mean value

References

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

National Assessment Governing Board, U.S. Department of Education. (2008) *Science framework for the 2009 national assessment of educational progress*. Washington, DC: U.S. Government Printing Office.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

Appendix
Terms of Use Agreement
Properties of and Changes in Matter Elementary School Student Assessment

By using the AIM Properties of and Changes in Matter Student Assessment developed by Horizon Research, Inc. (HRI), you agree to abide by the stipulations below concerning use, test security, test administration, and citations.

Use of the Assessment

The Properties of and Changes in Matter Student assessment may be used to gauge growth in knowledge about a specific content area as a result of an intervention such as professional development, curriculum use, or mentoring. It may also be used to learn about the contribution of teacher knowledge to student knowledge and classroom instruction.

We ask that you abide generally by the standards put forward in the *Standards for Educational and Psychological Testing* (AERA/APA, 1999).

You may not use the assessment to evaluate individuals. Assessment results may not be associated with any high-stakes consequence such as tenure, pay, hiring, or grades. The assessments were not developed for making decisions/judgments about individuals.
You should also refrain from using these measures to publicly demonstrate teachers' ability or lack of ability in science, which may adversely affect willingness to participate in future studies.

IRB and/or District/School Study Approval

It is your responsibility to obtain proper IRB and/or the appropriate district/school approval for your study and to follow the necessary requirements for obtaining principal, teacher, parent, and/or student permission/approval to administer to the assessment(s).

Responsibilities to Teachers and Students

Your responsibilities to study participants will largely depend on the details of the IRB and/or district/school approval of your study. In most cases, completion of the assessment will be strictly voluntary. As such, participants should be informed of the voluntary nature of the study. Teachers should be assured that if their students' data are not anonymous, individual identities will be kept strictly confidential; i.e., an individual's score or responses will never be reported in association with his or her name or any other identifying information. To encourage a high response rate among teachers, it may be helpful to:

- Clearly explain what the data will be used for and why the data are important for your study;
- Explain that there are no high-stakes consequences associated with completing the assessment; and
- Offer teachers compensation for time spent outside of the regular school day completing the assessment.

Test Security

The AIM Properties of and Changes in Matter Student Assessment may NOT be shared without prior authorization from HRI. Anyone who administers the assessment must agree to:

- Refrain from using any non-released item in any presentation, paper, article, or other public forum. Items are expensive to develop and pilot, and we are attempting to keep our item pool secure.
- Refrain from distributing copies of any non-released item to individuals other than participants in your research project.
- Refrain from using the assessment, in original or in copied form, to provide test-taking practice or to enhance test-taking skills.
- Refrain from using test items, actual or similar, for discussion or review.

(HRI acknowledges that, in some cases, school administrators and IRBs may require that the test materials be reviewed prior to granting permission for study participants to take the test. Such a review is not considered a violation of this Test Security Policy as long as the other provisions of this policy are not violated.)

Citing AIM Assessments

In any writing in which data from HRI's AIM assessments are included, the following citation must be used:

The assessment was developed by the Assessing the Impact of the MSPs: K–8 Science (AIM) project at Horizon Research, Inc., funded by the National Science Foundation under grant number DUE-0928177. Any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation or Horizon Research, Inc.

By signing below, I acknowledge that I have read the user manual, and I agree to abide by terms of use described above.

____

Printed Name Signature Date

Address: ____

Street City State Zip code

Phone number (including area code): ______________________________________

Your email address: ____