# AIM User Manual
# Evolution and Diversity
# Middle School Teacher Assessment

## Overview

The AIM Evolution and Diversity Middle School Teacher Assessment is a 30-item multiple-choice assessment developed for middle grades science teachers. 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 related single content area: Evolution and Diversity - differences and adaptations of organisms, preferential survival and relatedness of organisms.

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 Evolution and Diversity Teacher 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 areas. These pairs of assessments enable the study of relationships between teacher 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 elementary school and middle school levels.

[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.

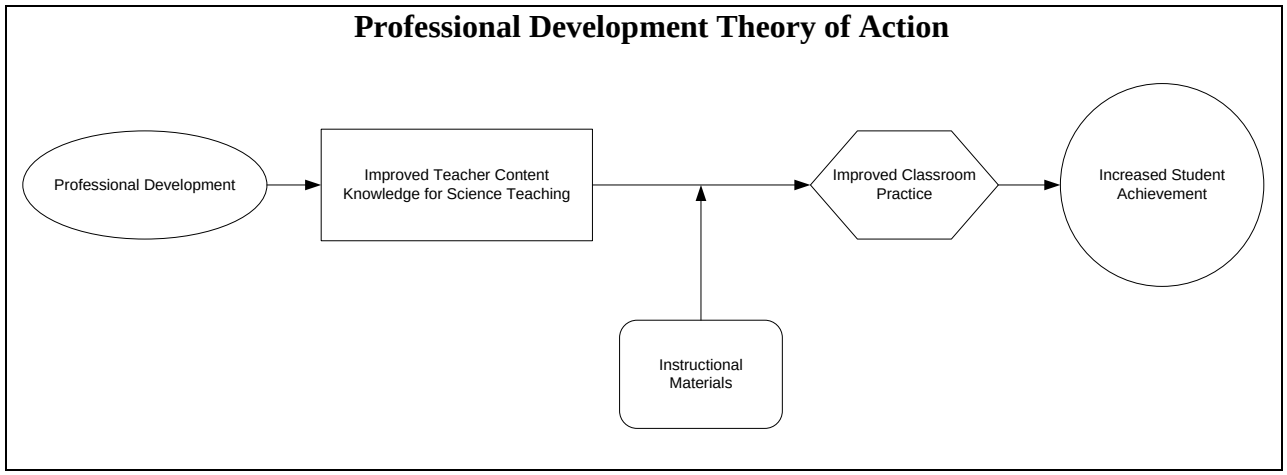


***Figure 1***

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 Evolution and Diversity Middle School Teacher 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).

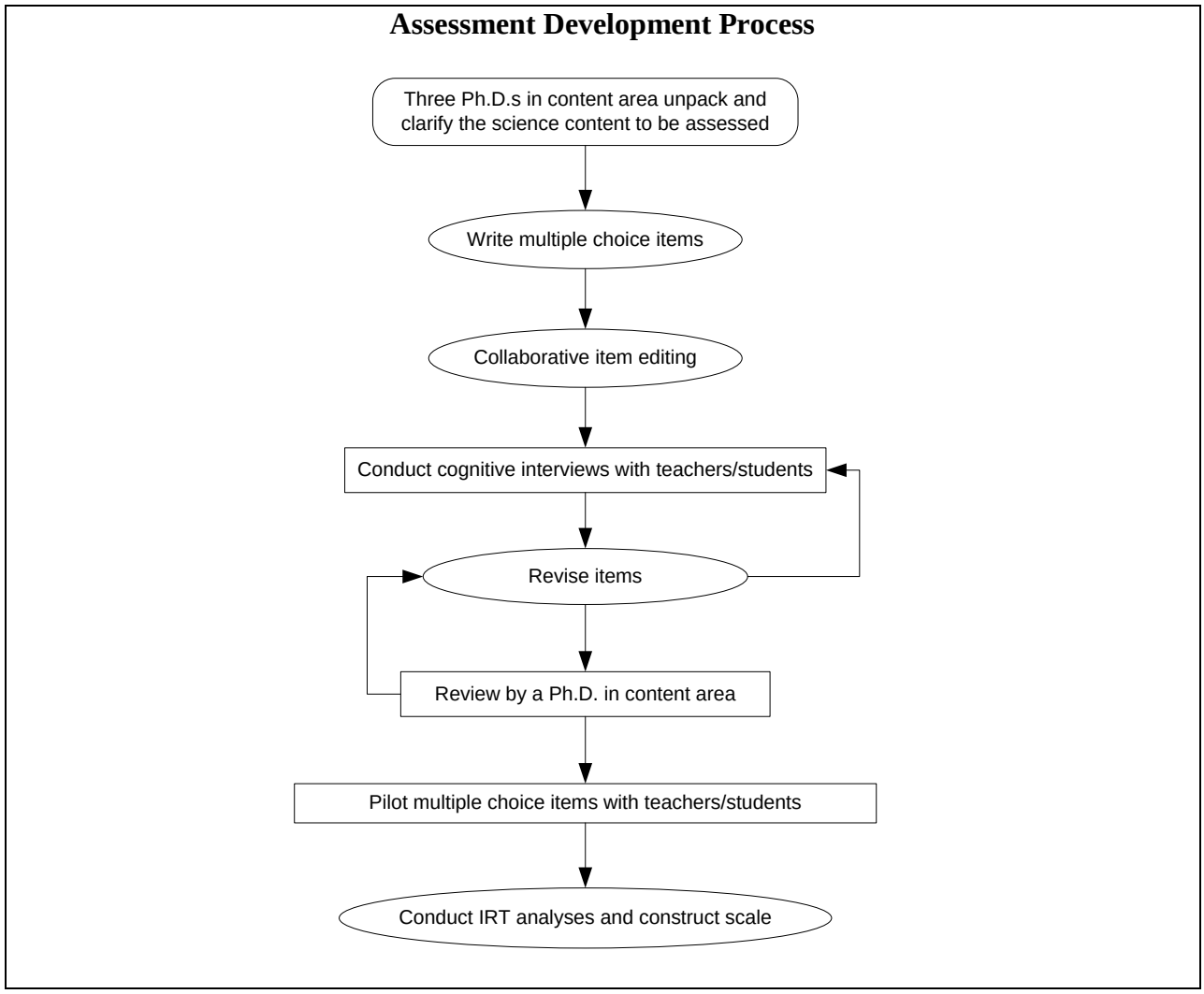


***Figure 2***

## Clarifying the Content Domain

Development began with identifying the target content for the evolution and diversity 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 a single strand in the *NAEP Framework*: differences and adaptations of organisms, preferential survival and relatedness of organisms. This process had three biologists/biology educators "unpack" the content into series of "sub-ideas" for middle school students. Additional sub-ideas that are important for teachers to understand in order to teach the middle school student ideas were also specified by the content experts. Both sets of ideas were considered in developing the middle school teacher assessment. The final description of the content domain is shown in Table 1.

## Table 1a
## Evolution and Diversity Content Domain

**Evolution and Diversity**. Differences and adaptations of organisms, preferential survival and relatedness of organisms

**Sub-ideas for students:**

- Different types of organisms (including plants and animals) have characteristics that enable them to survive and reproduce in different environments.
  - Different environments have different features that affect organisms' abilities to survive and reproduce. Some important features are climate, light level, soil nutrients, and the presence of other organisms.
  - Organisms that have characteristics that best meet the challenges of their environment are most likely to survive and reproduce. Therefore, the living organisms found in any particular environment are most likely to be those with characteristics that best meet the challenges in that environment.
  - A characteristic provides an advantage if it usually allows the number of individuals of that type of organism to increase over time; a characteristic provides a disadvantage if it usually causes the number of individuals of that type to decrease.
  - A characteristic or set of characteristics that provides an advantage in one environment is likely to be different than one that provides an advantage in other environments.
  - As environments change over time, the most successful adaptations will change, too.
  - Different types of organisms may meet the same environmental challenges in different ways. (For instance in a dry environment, animals may have leathery skin and plants may have a waxy covering, both of which decrease water loss.)
- Individuals of the same type differ in their characteristics, and sometimes the differences give individuals an advantage in surviving and reproducing.
  - There is variation in characteristics of organisms, even of the same type.
  - In a given environment, organisms with variations of a characteristic are likely to differ in their ability to survive and reproduce because some variations work better than others in that environment.
- Individual organisms with certain characteristics in particular environments are more likely than others to survive and have offspring.
  - Organisms of the same type in a particular environment have variation in many of their characteristics. No two individuals are exactly alike.
  - Some characteristics are more likely than others to lead to survival and reproduction in a particular environment.
  - Organisms inherit their characteristics from their parents. An individual with a characteristic that provides an advantage for survival and reproduction is more likely to have offspring that also have that advantageous characteristic. As these individuals survive and reproduce, their advantageous characteristics become more common in the population of organisms.
- When an environment changes, the advantage or disadvantage of characteristics can change.
  - Environments are not stable; they change from time to time. Environmental changes can be physical (e.g., temperature, rock slide), chemical (e.g., soil pH), biological (e.g., predators immigrating, competitors introduced) or a combination of these.
  - A characteristic that provides an advantage in one environment may provide less or more of an advantage, provide a disadvantage, or be neutral in a different environment.
- Extinction of a type of organism occurs when the environment changes and the characteristics of a type of organism are insufficient to allow its survival. Fossils indicate that many organisms that lived long ago are extinct. Extinction of organisms is common; most of the types of organisms that have lived on the Earth no longer exist.
  - Although organisms have characteristics that have variation, those variations are not always sufficient to meet the challenges of the environment. Individuals may die before reproducing or may not reproduce enough to replace all individuals who die.
  - If all members of a type of organism die, the type of organism becomes extinct.
  - Environmental changes that lead to the extinction of a type of organism often do so over time, such as through increased competition for limited resources, increased incidence of disease or parasites, or reduced ability to escape predators. Large-scale "catastrophic" changes in an environment may also produce extinctions.
  - There were many types of organisms in the past that were different than the ones that are alive today. We know this because we have a fossil record of some of the types of organisms that inhabited Earth in the past.
  - Fossils can provide scientists with records of organisms' structures that help identify the organisms and provide information about where and how they lived.
  - Types of organism can exist for long periods of time (many thousands to millions of years) before they become extinct. Becoming extinct does not mean a type of organism was unsuccessful; any type of organism may become extinct if its characteristics no longer provide an advantage in its environment.
- Similarities among organisms are found in anatomical characteristics, which can be used to infer the degree of relatedness among organisms.
  - Organisms of the same type have similar anatomical characteristics. Some of these characteristics can be seen with the naked eye, and some are microscopic.
  - The more closely related organisms are, the more similar their anatomy is likely to be; however, similar anatomy on its own does not prove relatedness.
- In classifying organisms, biologists consider details of internal and external structures to be more important than behavior or general appearance.
  - Grouping organisms according to their relatedness is called *classification*.
  - In judging similarity or differences of organisms, internal and external anatomical features should be examined.
  - Behavior and general appearance can be used to consider relatedness of organisms, but they are less reliable than specific anatomical features. Similar behaviors (e.g., flight, nest-building, earth-digging) are more likely than anatomical features to be shared by diverse organisms.
- Similar characteristics are not always an indication of relatedness because sometimes unrelated organisms may evolve similar adaptations to similar environmental constraints in similar ways. If the similar characteristics evolved from different structures that have different developmental origins, the process of evolving similar characteristics is called *convergent evolution.*

**Table 1b**
**Evolution and Diversity Content Domain**

| **Evolution and Diversity**. Differences and adaptations of organisms, preferential survival and relatedness of organisms |
|---|
| **Sub-ideas for teachers:**<br>• A species' ability to adapt is limited by the range of heritable variation present for each of its characteristics. Sources of heritable variation include mutations (changes to the genetic code) and sexual recombination.<br>• Mutations are random; new characteristics do not come about simply because the environment presents a need for them.<br>• Evolution often occurs through *natural selection*, which is the survival and reproduction of organisms that carry variations of characteristics that provide an advantage in a given environment. The environment constrains what variations provide an advantage and, therefore, which variations are most likely to be passed from generation to generation.<br>• Although natural selection promotes the survival of organisms with traits that are advantageous for a given environment, it does not guarantee survival. Some organisms with advantageous traits will not survive and/or reproduce. In other words, natural selection acts at the level of the individual, not the characteristic.<br>• Evolution may also occur due to *genetic drift*, which is the random, non-selection-based survival and reproduction of organisms. In other words, some variations of characteristics may persist and others may disappear through chance events.<br>• Evolution is a change in the frequency of heritable traits within a population or species over time. Evolution may result in new species developing, but does not always.<br>• Relatedness refers to how recently there has been a common ancestor for different types of organisms (i.e., species).<br>• Differences in characteristics arise from heritable variation among individuals of a type of organism. These differences arise from a variety of sources including mutation and sexual recombination.<br>• The fossil record is incomplete. There were likely many more types of organisms inhabiting Earth in the past than are preserved in the fossil record.<br>• Extinction provides opportunities for other organisms to evolve to fill niches. For example, extinction of the dinosaurs provided mammals, which were small rodent-like organisms while dinosaurs dominated the Earth, an opportunity to evolve into the various forms with which we are familiar today.<br>• Populations living in different environments can evolve different adaptations. When populations differ enough from one another, they are referred to as species.<br>• Comparisons of genetic material are another important tool in determining relatedness. The more closely related two species are (i.e., the more recently they had a common ancestor), the more genetic material they will have in common. (For example, scientists have determined that the genetic material of humans and chimpanzees is 96 percent the same, and there are fewer similarities between the genetic material of humans and other species.) |

## Types of Teacher Assessment Items

The sections that follow discuss the three types of teacher multiple-choice items included in the assessment:

1. knowledge of science content (Level 1 items);
2. using content knowledge to analyze/diagnose student thinking (Level 2 items); and
3. using content knowledge to make instructional decisions (Level 3 items).

The example items below are not included in the AIM assessment, but are shown here to illustrate item features. These example items may be flawed and are not intended to be used in any assessments.

### *Knowledge of science content*

All of the AIM items for teachers assess knowledge of science content, but the most basic type of question attempts to isolate disciplinary content knowledge from a teachers' ability to apply that

knowledge in making instructional decisions. An example "Level 1" item is shown in Figure 3 (correct answer is C).

**Level 1 Item**

A teacher reads an article about the discovery of an animal species that became extinct millions of years before dinosaurs lived. Based on its characteristics, a group of scientists thinks that some offspring of this species gave rise to modern birds and other offspring gave rise to dinosaurs.

If these scientists are correct, what is true about the relatedness of modern birds and dinosaurs?

A. Modern birds and dinosaurs are not related because they are different species.
B. Modern birds and dinosaurs are closely related because they shared a common ancestor.
**C. Modern birds and dinosaurs are distantly related because their common ancestor lived long before them.**
D. Modern birds and dinosaurs are related but it is not possible to tell how closely because they lived at different times.

***Figure 3***

This item illustrates some features common to all AIM teacher assessment items. As mentioned previously, all of the items are multiple-choice. In addition, all items include only four choices, and answer choices are never worded as "all of the above" or "none of the above." Multiple correct answers, such as "A and B, but not C," are also not used. Perhaps most importantly, all of the items are set in an instructional context. The intent in using these contexts was two-fold. First, we wanted teachers to feel like they were taking a test that was written for them, as opposed to, for example, a test constructed for undergraduates. The second goal was for teachers to recognize in the items the kind of work they do every day, making it more likely that they would intellectually engage with the items.

***Using science content knowledge to analyze/diagnose student thinking***

"Level 2" items require teachers to apply their content knowledge in analyzing or diagnosing a sample of student thinking. Figure 4 shows an illustrative item (correct answer is C).

**Level 2 Item**

A student writes the following in her biology journal.

"I learned today that badgers dig to make their homes under the ground. Worms do that, too. Cats are different because they live above ground and often climb trees. Even though both badgers and cats have bones and eyes and fur, badgers must be more closely related to worms than they are to cats because badgers and worms both dig in the earth."

What important idea regarding organism relatedness is missing in this passage?

A. Similar behaviors result in similar anatomy for unrelated types of organisms.
B. Both internal and external structures should be considered when determining relatedness.
**C. Similar behaviors are more likely than anatomical features to be shared by unrelated types of organisms.**
D. No important idea regarding organism relatedness is missing in this student's passage.

***Figure 4***

Certainly a teacher must understand the science content in order to select the correct answer choice for Level 2 items. However, additional analysis of the question is required because more than one of the choices includes a correct science statement, unlike the Level 1 item in Figure 3. In Figure 4, the statements in choices B, and C are correct in terms of the science, but only C applies to what the students said. This feature is present in all Level 2 items and makes the cognitive load of these items higher than that of Level 1; teachers must evaluate the students' thinking in relation to the science context in order to determine which answer choice is correct.

***Using content knowledge to make instructional decisions***

"Level 3" items ask teachers to apply their content knowledge in choosing among instructional moves. A sample Level 3 item is shown in Figure 5 (correct answer is C).

**Level 3 Item**

While reading articles from a science education journal, a teacher comes across the following research observation:

**Observation**

In a certain population of rats, it is usually advantageous to have brown fur because it provides better camouflage. Researchers take a sample of the population and notice that there are fewer rats with brown fur than with grey fur. The researchers also note that there have been no major changes to the rats' environment.

Are the researchers' observations consistent with the theory of natural selection?

A. No, the findings are inconsistent with the idea that natural selection guarantees the survival of individuals with an advantageous trait.
B. No, the findings are inconsistent with the idea that natural selection results in more variation of traits within a population.
**C. Yes, the findings are consistent with the idea that natural selection does not guarantee the survival of individuals with an advantageous trait.**
D. Yes, the findings are consistent with the idea that natural selection acts only on those individuals that are the strongest and not on individuals with advantageous traits.

***Figure 5***

Level 3 items have the highest cognitive load; teachers must evaluate the science content, the student's thinking in relation to the science content, and then evaluate each instructional choice. As with Level 2 items, more than one answer choice contains statements consistent with a correct interpretation of the science content, but only one has a correct science statement *and* is relevant to the instructional context. Although the cognitive load of Level 3 items is demanding, it is a small fraction of the demand placed on a teacher managing the learning of a classroom of students.

**Item Development**

HRI staff drafted items individually then met to edit them collaboratively. As the pool of items grew, we began recruiting middle school teachers for telephone cognitive interviews. We interviewed at least three teachers on each item in the pool using the interview protocol shown in Figure 7. After a round of interviews, HRI staff met to discuss teachers' feedback. If substantive edits were made to an item, we interviewed additional teachers 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 Teacher Assessment Items**
**Cognitive Interview Protocol**

**Prologue Script:**

Thank you for agreeing to let us interview you. As we explained in the email, we are developing a test for middle school science teachers, and we need your help to refine the test questions. I don't expect you to get all of the answers right. The point is to help us write a good test, not to test what you do or don't know. Do you have any questions before we get started? Remember that all of your answers are confidential. If you decide you would like to stop at any point, just say so.

**Procedure:**

- Ask teacher to read aloud and "think aloud" as they read the questions and answer choices, if they are comfortable doing so.

- 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 teachers?
  6. Do you have any other comments on the item?

***Figure 6***

The cognitive interviews revealed distinct patterns of errors in teacher responses to the Level 2 items (using content knowledge to analyze/diagnose student thinking). Some teachers chose an answer that included student thinking they were familiar with, whether or not it represented the thinking of the student in the item. Others chose a statement that was correct in terms of the science, but not in relation to the student's thinking.

Interviews also suggested some common errors teachers make with Level 3 items (using content knowledge to make instructional decisions). First, they often saw more than one of the instructional choices (including the correct one) as equally good, particularly when the item requires teachers to evaluate which question should be asked next. When the choices are about actual activities, teachers sometimes get bogged down in the details of the choices. For instance, they may rule out a choice that requires particular equipment because they do not have access to such equipment, regardless of whether the activity would help move the student's thinking forward. Finally, teacher beliefs about effective instruction may get in the way, even when they seem to understand the content targeted by the item. For example, teachers often choose a hands-on activity, even if it does not address the student's thinking.

**Pilot**

We selected 39 items to pilot with 443 teachers recruited from mailing lists of middle grades teachers across the country. The pilot was administered via the Internet. Approximately 20 percent of the sample was comprised of teachers who indicated they had taught the content at the

high school level to ensure that some respondents would be at the upper end of the knowledge spectrum.

**Table 2**
**Characteristics of the Pilot Test Sample (N = 443)**

| | Percent of Teachers |
|---|---|
| **Grade Levels Taught in 2010–11**[†] | |
| 6th grade | 21 |
| 7th grade | 36 |
| 8th grade | 42 |
| 9th grade | 21 |
| 10th grade | 22 |
| 11th grade | 21 |
| 12th grade | 19 |
| **Taken a college-level biology course beyond the introductory course** | |
| Yes | 81 |
| No | 19 |
| **Gender** | |
| Female | 66 |
| Male | 34 |
| **Race/Ethnicity**[†] | |
| American Indian or Alaskan Native | 2 |
| Asian | 2 |
| Black or African American | 2 |
| Hispanic or Latino | 2 |
| Native Hawaiian or Other Pacific Islander | 0 |
| White | 94 |

[†] Percentages may add up to more than 100 as teachers 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, 30 items were selected for the final form. The distribution of items by sub-idea is shown in Table 3. The number of items totals to more than 30 because one item may address more than one sub-idea. There are fewer sub-ideas in Table 3 than in the content unpacking (see Table 1), as limiting the assessment to a total of 30 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. Evolution often occurs through *natural selection*, which is the survival and reproduction of organisms that carry variations of characteristics that provide an advantage in a given environment. The environment constrains what variations provide an advantage and, therefore, which variations are most likely to be passed from generation to generation. | 2 |
| B. Individual organisms with certain characteristics in particular environments are more likely than others to survive and have offspring. | 1 |
| C. Some characteristics are more likely than others to lead to survival and reproduction in a particular environment. | 1 |
| D. Organisms inherit their characteristics from their parents. An individual with a characteristic that provides an advantage for survival and reproduction is more likely to have offspring that also have that advantageous characteristic. As these individuals survive and reproduce, their advantageous characteristics become more common in the population of organisms. | 3 |
| E. When an environment changes, the advantage or disadvantage of characteristics can change. | 2 |
| F. Environments are not stable; they change from time to time. Environmental changes can be physical (e.g., temperature, rock slide), chemical (e.g., soil pH), biological (e.g., predators immigrating, competitors introduced) or a combination of these. | 1 |
| G. A characteristic that provides an advantage in one environment may provide less or more of an advantage, provide a disadvantage, or be neutral in a different environment. | 1 |
| H. Extinction of a type of organism occurs when the environment changes and the characteristics of a type of organism are insufficient to allow its survival. Fossils indicate that many organisms that lived long ago are extinct. Extinction of organisms is common; most of the types of organisms that have lived on the Earth no longer exist. | 3 |
| I. Although organisms have characteristics that have variation, those variations are not always sufficient to meet the challenges of the environment. Individuals may die before reproducing or may not reproduce enough to replace all individuals who die. | 2 |
| J. If all members of a type of organism die, the type of organism becomes extinct. | 1 |
| K. Fossils can provide scientists with records of organisms' structures that help identify the organisms and provide information about where and how they lived. | 1 |
| L. Types of organism can exist for long periods of time (many thousands to millions of years) before they become extinct. Becoming extinct does not mean a type of organism was unsuccessful; any type of organism may become extinct if its characteristics no longer provide an advantage in its environment. | 3 |
| M. Similarities among organisms are found in anatomical characteristics, which can be used to infer the degree of relatedness among organisms. | 1 |
| N. The more closely related organisms are, the more similar their anatomy is likely to be; however, similar anatomy on its own does not prove relatedness. | 1 |
| O. In classifying organisms, biologists consider details of internal and external structures to be more important than behavior or general appearance. | 1 |
| P. Grouping organisms according to their relatedness is called *classification*. | 1 |
| Q. Behavior and general appearance can be used to consider relatedness of organisms, but they are less reliable than specific anatomical features. Similar behaviors (e.g., flight, nest-building, earth-digging) are more likely than anatomical features to be shared by diverse organisms. | 1 |
| R. Similar characteristics are not always an indication of relatedness because sometimes unrelated organisms may evolve similar adaptations to similar environmental constraints in similar ways. If the similar characteristics evolved from different structures that have different developmental origins, the process of evolving similar characteristics is called *convergent evolution*. | 2 |
| S. Relatedness refers to how recently there has been a common ancestor for different types of organisms (i.e., species). | 2 |
| T. Differences in characteristics arise from heritable variation among individuals of a type of organism. These differences arise from a variety of sources including mutation and sexual recombination. | 1 |
| U. The fossil record is incomplete. There were likely many more types of organisms inhabiting Earth in the past than are preserved in the fossil record. | 1 |
| V. Populations living in different environments can evolve different adaptations. When populations differ enough from one another, they are referred to as species. | 1 |

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** | **K** | **L** | **M** | **N** | **O** | **P** | **Q** | **R** | **S** | **T** | **U** | **V** |
| 1 | D | | | | | | | | | P | | | | | | | | | | | | | |
| 2 | D | | | | | | | | | | | | | | | | | | | P | | | |
| 3 | D | S | | | P | | | | | | | | | | | | | | | | | | |
| 4 | B | | | | | | | | | | | | | | | | | | | | P | | |
| 5 | A | | | | | | | | | | | | | | P | | | | | | | | |
| 6 | A | | | | | | | | P | | | | | | | | | | | | | | |
| 7 | B | | | | | P | | | | | | | | | | | | | | | | | |
| 8 | D | | | | | | | | | | P | | | | | | | | | | | | |
| 9 | B | | | | | | | | | | | | P | | | | | | | | | | |
| 10 | C | | | | | | | | | | | | | | | | | | P | | | | |
| 11 | D | | | | | | | | | | | | | | | | P | | | | | | |
| 12 | A | | | | | | | | | | | | P | | | | | | | | | | |
| 13 | A | | P | | | | | | | | | | | | | | | | | | | | |
| 14 | D | | | | | | | | | | | | | | | P | | | | | | | |
| 15 | D | | | | | | | | | | | P | | | | | | | | | | | |
| 16 | A | | | | | | | | | | | | P | | | | | | | | | | |
| 17 | B | | | | | | | | | | | | | | | | | | P | | | | |
| 18 | D | | | | | | | | P | | | | | | | | | | | | | | |
| 19 | B | | | | | | P | | | | | | | | | | | | | | | | |
| 20 | A | | | | | | | | | | | | | | | | | P | | | | | |
| 21 | B | | | | | | | P | | | | | | | | | | | | | | | |
| 22 | B | | | | | | | | | P | | | | | | | | | | | | | |
| 23 | C | | | P | | | | | | | | | | | | | | | | | | | |
| 24 | C | | | | | P | | | | | | | | | | | | | | | | | |
| 25 | B | | | | | | | | | | | | | P | | | | | | S | | | |
| 26 | C | | | | P | | | | | | | | | | | | | | | | | | |
| 27 | D | | | | | | | | P | | | | | | | | | | | | | | |
| 28 | A | | | | | | | | | | | | | | | | | | | | | P | |
| 29 | A | S | | | P | | | | | | | | | | | | | | | | | | |
| 30 | C | | | | | | | | | | | | | | | | | | | | | | P |
| **Primary:** | | 0 | 1 | 1 | 3 | 2 | 1 | 1 | 3 | 2 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| **Secondary:** | | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| **Total:** | | 2 | 1 | 1 | 3 | 2 | 1 | 1 | 3 | 2 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 |

**Validity**
Three lines of evidence support the argument that the assessment is a valid measure of teachers' knowledge of these evolution and diversity ideas. First, cognitive interviews with teachers established that teachers interpret the items as intended and that teachers must use their knowledge of content to answer the items correctly. Second, a content expert (individual with a Ph.D. in biology) 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 for teaching about evolution and diversity."

**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 0.85; 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**
The pilot was not a timed administration. As such, there is no information about speededness.

## Using the Assessment

The AIM Evolution and Diversity Teacher 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 Evolution and Diversity Teacher Assessment yields a score for each individual. However, the assessment is not valid for making judgments about individuals based on those scores. For instance, evaluating teacher performance 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**
As described above, the pilot test was administered on-line and was not timed. Although there is no evidence of speededness, it is recommended that at least 30 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 | 12 |
| 2 | 20 |
| 3 | 25 |
| 4 | 28 |
| 5 | 31 |
| 6 | 34 |
| 7 | 36 |
| 8 | 39 |
| 9 | 41 |
| 10 | 43 |
| 11 | 45 |
| 12 | 46 |
| 13 | 48 |
| 14 | 50 |
| 15 | 51 |
| 16 | 53 |
| 17 | 55 |
| 18 | 56 |
| 19 | 58 |
| 20 | 60 |
| 21 | 62 |
| 22 | 64 |
| 23† | 66† |
| 24 | 68 |
| 25 | 70 |
| 26 | 73 |
| 27 | 76 |
| 28 | 81 |
| 29 | 88 |
| 30 | 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
# Evolution and Diversity Middle School Teacher Assessment

By using the AIM Evolution and Diversity Teacher 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 Evolution and Diversity Teacher 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 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 Evolution and Diversity Teacher 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: _____