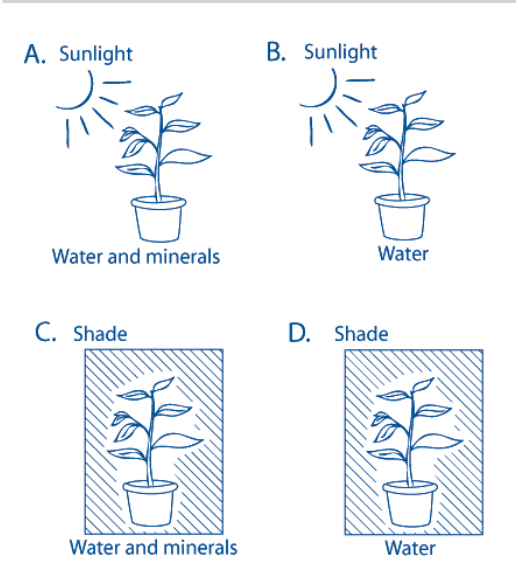
2061 today

Science Literacy for a Changing Future

Winter 2007 ▸ Volume 17, Number 1

Getting Assessment Right

Rigorous Criteria, Student Data Inform Project 2061's Assessment Design



With state testing in science under the federal No Child Left Behind (NCLB) law due to begin in the 2007–2008 academic year, assessments that accurately measure students' achievement of science standards are a high priority. Yet there is growing concern among educators and the public that today's assessments are not up to the task. According to a recent study from the American Federation of Teachers, only 11 states have strong content standards and tests that are aligned to them for reading, mathematics, and science.

But what are the features of effective standards-based assessment? And how can test questions be improved? To help K–12 science educators answer these questions, Project 2061's latest assessment work combines the analysis of test items using research-based criteria with the use of student data to improve each item's effectiveness.

With funding from the National Science Foundation, Project 2061 is building an online collection of high-quality middle school and early high school science assessment items that are linked to important ideas in AAAS's *Benchmarks for Science Literacy* and the National Research Council's *National Science Education Standards*. Useful for researchers, teachers, and textbook and test developers, the mostly multiple-choice items are designed to measure as precisely as possible what students do and do not know about the ideas and skills that are targeted in the national standards. The online collection will supplement the test items with a wealth of related assessment resources:

- *Clarifications* of each key science idea that pinpoint what students are expected to know;
- *Common student misconceptions* identified by research and useful for designing incorrect answer choices (distractors); and
- *Assessment maps* that show how key ideas build toward student understanding.

Users will be able to search by topic, by national benchmark or standard, and by keywords and phrases from state standards.

Testing the Tests

Project 2061's approach to developing assessment items aligned to standards involves multiple rounds of analysis by staff and outside experts, along with student interviews and pilot tests. Using a set of research-based criteria, Project 2061's research team first examines the alignment of each item to the ideas in a science standard. They then consider other features of the item that might make it difficult to interpret students' responses.

When an item is well designed, students should choose the correct answer only when they know the targeted idea and they should choose an incorrect answer only when they do not know the idea. Students should be able to demonstrate their knowledge of the science without being tripped up by confusing

Continued on page 4

WHAT'S INSIDE

Project 2061
AAAS

Jo Ellen Roseman

"The nation is in an excellent position to identify a uniform and coherent set of science standards."

FROM THE
Director

Making Tests Meaningful

Testing students in science under the federal No Child Left Behind law is just around the corner, due to begin in the 2007–2008 school year. While the goal of the new requirement is admirable, today's large-scale science tests simply aren't that good. Too often, as Project 2061 has found in its studies of assessment, the questions in such tests are confusing or not well aligned to the key science ideas and skills that students are expected to learn (see "Getting Assessment Right" in this issue).

For standards-based reform to succeed, it is important to create and use tests that are fair and accurate measures of student knowledge. But our science tests will only provide meaningful information if they are truly aligned to important science ideas, such as those in national benchmarks and standards. Earlier this year, AAAS applauded a new bipartisan proposal for voluntary nationwide standards in science and mathematics. Given the varying quality of current state science standards, The Standards to Provide Educational Achievement for Kids (SPEAK) Act, co-authored by Sen. Chris Dodd (D-CT) and Rep. Vernon Ehlers (R-MI), is a significant first step.

The good news is that the nation is in an excellent position to identify a uniform and coherent set of science standards. AAAS's statement on the new proposal recommended that planners begin with existing guidelines like Project 2061's *Benchmarks for Science Literacy* and the National Research Council's *National Science Education Standards* (*NSES*). These standards documents have provided a model for the nation's best state-level science standards. Of further help is the 2009 NAEP Science Framework, which has already done the work of synthesizing and updating the physical, life, and earth science standards found in *Benchmarks* and *NSES*.

Under the SPEAK Act, states that choose to adopt the new standards would then receive federal funds to implement them—for professional development, assessment, curriculum materials, and so on. With high-quality, nationwide standards in place, educators could pool resources to develop tests and classroom-based assessments that measure student progress toward the specific ideas and skills contained in those standards. More importantly, an agreed upon set of science standards would provide focus to curriculum materials developers and teacher educators. If we want to help all students develop a coherent and scientifically accurate picture of how the world works, we very much need to focus finite resources on teaching the most important science ideas.

Jo Ellen Roseman

To contact Project 2061 staff, visit www.project2061.org/about/contact.htm.

AAAS Resources for Educators

Atlas of Science Literacy, Volume 2

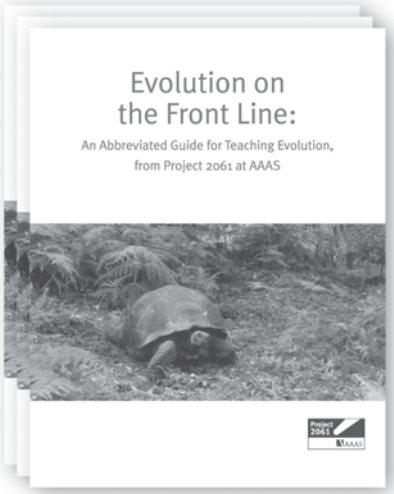
Reserve your copy of *Atlas of Science Literacy, Volume 2*, scheduled for release in March 2007. Completing the work begun in the first volume of the popular *Atlas of Science Literacy*, the new *Atlas 2* maps out what all students should learn as they move from kindergarten through 12th grade. *Atlas 2* features all new maps for more than 40 essential topics, including Weather and Climate, Human Development, and Explaining Evolution.

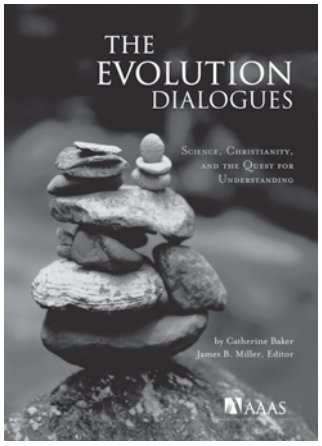
- Order *Atlas of Science Literacy, Volume 2* before March 31, 2007, at a special pre-publication price of $53.95 each (regular price is $59.95 each). *Atlas 1* and *Atlas 2* are also available as a two-volume set for $99.95. Call Project 2061 at 1-888-737-2061 (mention promotional code AAAS07) or go to **www.project2061.org/publications.order.htm**.

Evolution Guide in Print, Online

Today's students need a solid understanding of how life on Earth evolved if they are to become science-literate adults. AAAS Project 2061's *Evolution on the Front Line: An Abbreviated Guide for Teaching Evolution*, first introduced at the 2006 AAAS Annual Meeting, was developed to help educators convey core evolution concepts to their students. Excerpts from Project 2061 tools identify what students need to learn and show how students build over time a coherent understanding of biological evolution, natural selection, and the nature of science. Also featured are recommended trade books and suggestions from AAAS for how to respond to some of the most frequently asked questions about teaching evolution.

- To order print copies of the guide, call 1-888-737-2061 or visit **www.project2061.org/publications/order.htm**. $3.95 each; orders of 10 or more copies, $3.00 each. Prices include shipping and handling to domestic addresses.
- To read the guide free online, visit **www.project2061.org/evolutionguide**.





Bridging Science and Religion

A new book from AAAS's Dialogue on Science, Ethics, and Religion (DoSER) offers a thoughtful look at both the development of evolutionary theory from Darwin's time to the present and the diversity of Christian responses to the theory. *The Evolution Dialogues: Science, Christianity, and the Quest for Understanding*, written by Catherine Baker and edited by James B. Miller, addresses misunderstandings about what biological evolution is, what science is, and what views people of faith have applied to their interpretations of the science. Well received by biology teachers and teachers of adult religious education classes, the book has been supplemented by a study guide for use in group settings.

- For more information about *The Evolution Dialogues*, including ordering information and access to the free online study guide, see **www.aaas.org/spp/dser**.

Join Project 2061 at NARST

Hear Project 2061 staff discuss their assessment studies at the National Association for Research in Science Teaching (NARST) Annual Conference in New Orleans, LA, April 15–18, 2007 (see **www.narst.org/conference**).

- "Assessment Linked to Science Learning Goals: Probing Student Thinking Through Assessment" (Symposium)
  Led by George DeBoer
- "Determining the Appropriateness of Terminology in Content-Aligned Assessment of Middle School Students: Examples from Plate Tectonics" (Poster)
  Paula Wilson and George DeBoer
- "Probing Middle School Students' Understanding of Ideas about Chemistry through Content-Aligned Assessment" (Poster)
  Cari Herrmann Abell and George DeBoer
- "Assessing Students' Understanding of 'Controlling Variables'" (Poster)
  Arhonda Gogos and George DeBoer

Getting Assessment Right

Continued from page 1

language, inaccurate information, unclear diagrams and graphs, or contexts that are unfamiliar or unnecessarily complex. Similarly, students who do not know the science should not be able to answer correctly simply because the wrong answers are not plausible or are otherwise susceptible to test-taking strategies.

Students should be able to demonstrate their science knowledge without being tripped up by confusing language, unclear diagrams, or contexts that are unnecessarily complex.

What Students Think

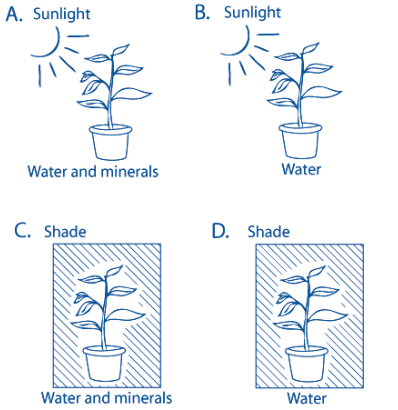
Researchers go a step further by combining their detailed analyses of a test item with information about how and why students actually select their answer to a question. Through one-on-one interviews with students and pilot tests of items in which students' answers are compared to the written explanations they give for their answers, Project 2061 is gaining valuable insights into student thinking about the targeted ideas and about the test items themselves.

"The student feedback we're getting from the pilot tests is proving essential to the item development process. It gives us information about misconceptions students have, terminology they are unfamiliar with, and phrasing they find confusing. We use this information to make the items more valid measures of student learning," said George DeBoer, deputy director of Project 2061. "Beyond that, the feedback helps us evaluate the appropriateness of the learning goals for middle school students, and in this way a close examination of the assessment items is also a useful tool in the refinement of those learning goals."

To find out if an assessment item is measuring the targeted idea or yielding unreliable information about what students know, students are asked the following questions in pilot tests:

1. Is there anything about this test question that was confusing? Explain.
2. Circle any words on the test question you don't understand or aren't familiar with.
3. Is answer choice A correct? Explain why.
4. Is answer choice B correct? Explain why.
5. Is answer choice C correct? Explain why.
6. Is answer choice D correct? Explain why.
7. Did you guess when you answered the test question?
8. Please suggest additional answer choices that could be used.
9. Was the picture or graph helpful? If there was no picture or graph, would you like to see one?
10. Have you studied this topic in school?
11. Have you learned about it somewhere else (TV, museum visit, etc.)? Where?

Analyzing the collected data gives researchers detailed information to use in revising individual questions and answer choices. (See box on page 5 for a look at what pilot tests revealed about a sample item.)



Putting It All Together

Project 2061 has already conducted 200 interviews with students in seven schools and administered pilot tests in 112 classrooms across five content areas: atoms and molecules, force and motion, flow of matter and energy, plate tectonics, and control of variables. The pilot tests were carried out in both urban and suburban middle and high schools serving a wide range of students, and more pilot tests are planned for 2007. Meanwhile, the research team is busy identifying the best test items developed so far and bringing them together with related resources for use in the online collection.

Once the collection is launched, teachers and researchers will be able to use the model items as well as Project 2061's criteria and analysis procedures to study and improve assessment in curriculum development projects, large-scale assessment programs, and classrooms. With high-quality resources aligned to standards, science educators will have a better chance of getting assessment right and supporting students' achievement of science literacy.

For more information about Project 2061's assessment work, visit www.project2061.org/assessment or contact deputy director George DeBoer at gdeboer@aaas.org or (202) 326-6624.

Pilot-Testing Items: What We Learn

Can teachers have confidence in a test question if students can get the right answer without really knowing the targeted science idea? To improve a test item's alignment to key science ideas and its validity as a measure of student learning, Project 2061 uses feedback from students gathered in pilot tests. The student data helps researchers identify and eliminate confusing or unfamiliar language in the item; refine the correct answer choice; and provide more plausible incorrect answer choices (distractors). Here's what pilot testing revealed about an item aligned to an idea about chemical reactions:

Targeted Idea

Substances may react chemically in characteristic ways with other substances to form new substances with different characteristic properties (based on *National Science Education Standards*, 5-8B:A2a)

Item

Which of the following is an example of a chemical reaction?
A. A piece of metal hammered into a tree.
B. A pot of water being heated and the water evaporates.
C. A spoonful of salt dissolving in a glass of water.
D. An iron railing developing an orange, powdery surface after standing in air.

Students Who Selected Each Answer Choice

	A (metal)	B (evaporation)	C (dissolving)	D (rusting)	Not sure	Total
Number of students	0	14	18	43	1	76
% of students	0	18.4	23.7	56.6	1.3	100

Results of Pilot Testing

- Only 5 of the 43 students who chose the correct answer D said that a new substance formed. Approximately half of the 43 students who chose D said they recognized it as an example of rusting or oxidation. Perhaps these students really do know that rusting is a chemical reaction that produces new substances with different properties, but they may also know rusting only as a specific instance of a chemical reaction without knowing that chemical reactions involve the formation of new substances.
- None of the students chose A, suggesting that hammering a piece of metal into a tree is not a plausible answer. Similar results were found during student interviews.
- A significant number of students (42.1%) chose either B or C. This supports other research that shows that students hold the misconception that phase change and/or dissolving are chemical reactions.

Suggested Revisions

- Replace answer choice A with a more plausible distractor, such as, "Sand being removed from sea water by filtration."
- Replace answer choice D with a reaction that students are not so familiar with, for example, "A white solid forming when two clear liquids are mixed together."

About AAAS and Project 2061

Publisher of the peer-reviewed journal *Science*, the American Association for the Advancement of Science (AAAS) is the largest general scientific organization in the world. Its education initiative, Project 2061, has been at the forefront of the K–12 reform movement

- Defining science literacy and promoting it as a goal for all Americans;
- Developing K–12 benchmarks for student learning in science, mathematics, and technology;
- Producing a wide range of innovative tools for educators—books, CD-ROMs, and online resources—to guide their reform efforts; and
- Conducting research on the design and use of curriculum materials, assessment, professional development, and other areas of science teaching and learning.

AAAS gratefully acknowledges the following Project 2061 supporters over the past 20 years: Carnegie Corporation of New York, Hewlett-Packard Company, John D. and Catherine T. MacArthur Foundation, Andrew W. Mellon Foundation, Noyce Foundation, David and Lucile Packard Foundation, Pew Charitable Trusts, Siemens Foundation, and the National Science Foundation.

For more information:
AAAS Project 2061
1200 New York Avenue, NW
Washington, DC 20005
Phone: 202-326-6666
Fax: 202-842-5196
E-mail: project2061@aaas.org
Web site: www.project2061.org

To order Project 2061 products:
Call 1-888-737-2061 or visit www.project2061.org/publications/order.htm

Better Instruments for Better Learning

An Interview With Sean Smith

Dr. Sean Smith is senior research associate at Horizon Research, Inc. (HRI), a private research firm located in Chapel Hill, North Carolina, that specializes in work related to science and mathematics education. Since he joined HRI in 1991, Dr. Smith has worked on a number of research and evaluation projects in the areas of curriculum materials development, professional development, and assessment. He has worked extensively with Project 2061's resources, most recently in the context of an instrument development project. Project 2061's Cathy Tramontana recently interviewed Dr. Smith via e-mail about his current work and his views on standards-based science education reform.

CT: Tell us a little about your background. What led you to get involved in science education research and evaluation?

SS: I'm a former high school chemistry and physics teacher. After four years in the classroom, I went back to graduate school in a Ph.D. program focused on science education. My experiences in graduate school contributed to my interest in research and evaluation, but an internship with HRI ultimately convinced me that I wanted to remain in this field. Research offers the opportunity for theory building. Evaluation work often occurs at the intersection of education theory and practice. I find both kinds of work very fulfilling.

"If states narrow and focus their standards, tests will follow suit, and teachers will eventually have more freedom to teach for understanding."

CT: Project 2061 is currently collaborating with you on the project "ATLAST (Assessing Teacher Learning About Science Teaching)." What are the project's goals and how will it help science teachers?

SS: At HRI, we do both research and evaluation in science and mathematics education. In both types of work, we're often confronted by the lack of well-established instruments in the field (assessments, surveys, observation protocols). We're convinced that a lack of instruments hinders the development of education theory. We're also struck by how much professional development for teachers operates on the largely untested theory that professional development ultimately leads to improved student learning. The theory is certainly logical, but has little empirical support in the literature.

The goal of ATLAST is to develop instruments that researchers can use to study the effects of professional development. We're developing measures of teacher opportunity to learn, teacher understanding of science, student opportunity to learn, and student understanding of science. Better insight into how professional development works should ultimately lead to better professional development experiences for teachers.

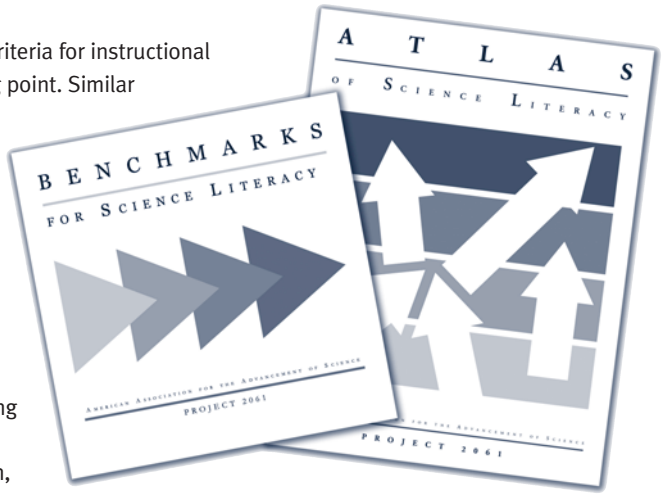
CT: How has Project 2061's approach to analyzing and developing student assessment items influenced your work and thinking?

SS: Project 2061's work impacts our work in a couple of ways. The first relates to *Benchmarks for Science Literacy* and *Atlas of Science Literacy*. We're developing our instruments in three areas of middle grades science: force and motion, plate tectonics, and flow of matter and energy in living systems. We spent months defining these content domains before we wrote our first assessment item. Both *Benchmarks* and *Atlas* were critical in this process. They helped us draw the boundaries around the content, establishing what is "fair game" for the assessments and the "opportunity to learn" instruments. Grades 9–12 benchmarks are clearly off limits for our student assessments. Grades K–2 and 3–5 benchmarks suggest content that we can reasonably expect middle grades students to understand. While our student assessment items do not test this earlier content directly, they often draw on the concepts. Because *Benchmarks* and *Atlas* helped us determine "what's in and what's out," our instruments are very tightly aligned to the content they are designed to measure.

Second, the Project 2061 assessment item criteria have strongly influenced our item construction. Project 2061 has reviewed over 150 of our student and teacher items. The criteria help us write items that are valid and reliable. We've been writing items for close to three years now, and at this point, the criteria are very much in our heads. As a result, our first drafts now are much closer to finished products, requiring many fewer revisions.

CT: In what ways are science curriculum materials being developed today different from those of the past? Why?

SS: My view on recently-developed science curriculum materials is probably skewed. Over the last six years, HRI has served as external evaluator on two major NSF-funded curriculum development projects. In each case, the project used Project 2061's *Benchmarks* and criteria for instructional materials as their starting point. Similar to HRI's ATLAST project, both groups spent months specifying and clarifying the content before developing the first activity. Both groups also took seriously the notion that activities and lessons must explicitly target one or more learning goals. That sounds like a painfully obvious criterion, but it's amazing how many currently available materials (traditional and reform-oriented) include activities that, while they address topics in a very engaging way, cannot be linked to a specific learning goal or are not well aligned to the learning goal they supposedly target.



That's the biggest difference I see in the new materials that I'm most familiar with. They are designed explicitly to go beyond engaging students to developing deep understanding of science concepts.

CT: What do you see as the major obstacles to the adoption and/or implementation of the new standards-based science curriculum materials?

SS: I think three obstacles stand in the way. First, these materials demand a different kind of instruction than I believe most science teachers are accustomed to. Successful implementation requires much more up front and ongoing support than that which traditionally accompanies textbook adoption. Such support is not cheap, but without it, I think the odds of teachers implementing the materials in a way that is consistent with the developers' intent are not very good.

Second, the materials themselves are expensive. They engage students with naturally occurring phenomena, and hands-on activities are in the forefront. While traditional textbooks certainly include hands-on investigations, I do not think these activities are as central to accomplishing learning goals as they are in the standards-based materials. The supplies and equipment associated with the standards-based materials are really not optional. As such, they may add a layer of cost that some states and schools simply will not bear, at least not until the research base on student learning clearly supports these materials.

Finally, state tests work against standards-based materials. In an attempt to develop deep understanding, standards-based materials have had to be more selective than traditional textbooks in the content they address. State standards, however, are still quite broad. At a topical level, traditional textbooks probably align better with state science tests.

CT: Given your work with Project 2061 as a collaborator and evaluator, how do you think the project can best make a difference in science education?

SS: Picking up on my last point, Project 2061 can and should continue to influence state standards. If states narrow and focus their standards, tests will follow suit, and teachers will eventually have more freedom to teach for understanding. Second, Project 2061's textbook evaluations have made a difference, but the battle for learning goal-focused instruction is far from won. The more ways that Project 2061 can make this point, the better.

For more information about Horizon Research, Inc., and its latest work, visit www.horizon-research.com.

Online Update

Are college students getting what they need from their science courses? Probably not, say Project 2061's Jo Ellen Roseman and Mary Koppal, writing in the new *Handbook of College Science Teaching* from NSTA Press. The chapter applies lessons learned at the K–12 level to science teaching and learning in a higher-education context. To learn more and read the chapter, visit **www.project2061.org/handbook.**

Moving?

Help *2061 Today* keep up with you! Send your change of address to project2061@aaas.org or call us at 202-326-6666. You may also visit **www.project2061.org/confirm** to update your address, switch from print to electronic delivery, be removed from the mailing list, or sign up for other FREE Project 2061 newsletters.

Announcements & Events

Register Now: 2007 *Atlas* Workshops

Project 2061 continues to offer "Using *Atlas of Science Literacy*" workshops. Join us for one of these popular events to learn how growth-of-understanding maps from *Atlas, Volume 1* and the new *Atlas 2* can help you to clarify standards and to improve curriculum, instruction, and assessment:

- October 16–18, 2007, Washington, DC

Scholarships are available and more workshops will be scheduled throughout 2007. To learn more and register, visit **www.project2061.org/workshops**.

New Projects for New Audiences

Thanks to AAAS's William T. Golden Fund for Program Innovation, Project 2061 has begun two new projects that will bring its standards-based resources to a wider range of formal and informal science educators:

- A "Consumer's Guide to Selecting High-Quality Instructional Materials" will help K–12 science and mathematics educators choose good instructional materials, whether in print or digital format, that are likely to contribute to student learning of important ideas. The guide will walk users through Project 2061's curriculum materials analysis procedure, provide illustrative examples, and include interactive tutorials on a companion CD-ROM.
- "Professional Development for Informal Science Institutions" will develop a three-day workshop customized to meet the needs of science centers, museums, zoos, and other informal institutions. The workshop will help educators at these institutions use Project 2061's tools to gain insights into what is known about effective science teaching and learning and how to interpret their institutions' resources through the lens of standards-based science education.

Workshop for NOAA

With new funding from the National Oceanic and Atmospheric Administration (NOAA), Project 2061 is customizing its *Atlas of Science Literacy* workshop to help NOAA develop a framework for climate and weather education. Drawing on *Atlas* and Project 2061's other standards-based resources, the spring 2007 workshop will address the nature and content of curriculum resources for the study of earth science. The workshop will also explore how federal agencies such as NOAA can ensure that their curriculum development efforts on climate and weather are of high quality and take advantage of national benchmarks and standards for K–12 science education. Invited participants will be able to draw on workshop resources to produce a framework to guide the development of programs, lessons, and other materials for the topic of climate and weather.

Introducing...

AAAS Project 2061 is pleased to welcome **Michael Anderson** as its new senior business analyst. He comes to Project 2061 from the U.S. Civilian Research & Development Foundation (CRDF) in Arlington, VA, where he was the grant financial manager for the Award Administration and Grants Assistance Program. Anderson holds a B.A. in economics and political science from The George Washington University.

Project 2061 AAAS

American Association for the Advancement of Science
1200 New York Avenue, NW
Washington, DC 20005

www.project2061.org

Address Correction Requested

Non-Profit Org.
U.S. Postage
PAID
Fredericksburg, VA
Permit No. 200