Project 2061
American Association for the Advancement of Science

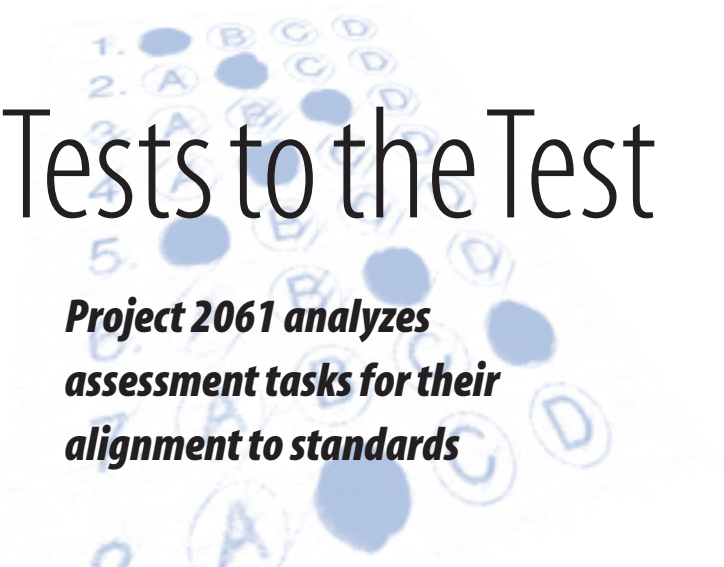
2061 today

Science Literacy for a Changing Future

Mathematics
Natural Sciences
Social Sciences
Technology

Spring/Summer 2001
Volume II, Number 1

Putting Tests to the Test



Project 2061 analyzes assessment tasks for their alignment to standards

Assessment is taking center stage as schools are held more accountable for student performance. Rewards and sanctions for schools are riding on test outcomes, and parents are using these scores to decide which schools are best for their children. Since what gets tested typically determines what gets taught, it is important that these assessments be carefully aligned to a school's science and mathematics standards for what students are intended to learn. Although there are some nationally recognized systems that survey tests for how well their content is distributed over a list of topics, there have been no useful guidelines for aligning assessments with specific learning goals—until now.

Project 2061 has received a $2.4 million grant from the National Science Foundation to develop new strategies and tools for evaluating the alignment of K–12 assessments in science and mathematics to national, state, and district standards and benchmarks. Using its experience in evaluating textbooks, Project 2061 is

- Developing an analysis procedure—and a computer utility for applying it—to judge alignment of assessment tasks to specific learning goals;
- Producing case studies that illustrate the use of the procedure to evaluate and revise existing assessment tasks and create new ones;
- Designing professional development workshops to help educators understand and apply the assessment procedure; and
- Building a collection of analysis profiles for the assessment tasks that are reviewed.

Full documentation and helpful background information on all aspects of this three-year project will be published in a book and web-connected CD-ROM. Project 2061 hopes that these resources will be useful to commercial developers and publishers of instructional and assessment materials, to districts and states that select and administer large-scale testing programs, and to classroom teachers who create or assemble their own quizzes or tests.

Andrew Ahlgren, associate director at Project 2061 and a leader of the assessment project, predicts that the work will result in marked improvement in assessments by influencing test developers and their customers in the schools. "On the one hand, we intend that our assessment analysis procedure will contribute to writing better tasks and tests," he says. "On the other hand, we hope to change educators' views on what to expect of assessment tasks and help them make better choices among what publishers offer them."

The Analysis Procedure in Detail

Project 2061's procedure evaluates an assessment task's potential to reveal whether students have attained a well-defined component of knowledge or acquired a particular skill. ("Task" is a general term for any selected-response, constructed-response, or performance-assessment "probe of student knowledge.")The analysis first identifies specific learning goals targeted by an assessment task, and then evaluates the likely effectiveness of the task in probing student achievement of those goals.

Project 2061's analysis does not focus merely on topic distribution—that is, on whether the

continued on page 2

Putting Tests to the Test *from page 1*

assessment tasks fit into goal topics and are distributed fairly among those topics. Instead, it probes more deeply to judge how well each individual task gets at the exact knowledge specified in a goal. Project 2061 considers an assessment task to be well aligned to a learning goal only if it both aims at the exact knowledge (fact, idea, or skill) specified in the goal and effectively probes student understanding of that specific goal.

Analysts first check whether the task has the minimum information required for analysis (such as a scoring guide and sample student responses) and then they create a manageable list of candidate goals that the task might target. Once the tasks are determined to be complete and the targeted learning goals are identified and clarified, analysts look closely at the following:

Content

Necessity: Is the specific knowledge that is to be assessed needed for all satisfactory responses?

Sufficiency: Is the knowledge enough by itself to make a satisfactory response or is something else also needed?

Likely Effectiveness

Comprehensibility: Are students likely to understand the task statement, diagrams, symbols, etc.?

Clarity of Expectations: Are students likely to understand what they are expected to do and what sort of response is considered satisfactory?

Appropriateness of Context: Is the task context appropriately familiar, engaging, and realistic to students?

Resistance to Test-Wiseness: Could students respond satisfactorily to the task by guessing or employing other general test-taking strategies?

Reliability of Scoring: Is the scoring guide accurate, clear, complete, and specific enough to yield consistent results for different reviewers?

Other Characteristics

Cost Effectiveness: Is what we learn about the student's knowledge worth the "cost" of the task in terms of time and effort?

Reusability: Can the task be given to students who have seen or heard about it before or can it be easily modified to do so?

Alternative Response Modes: Do students have an opportunity to demonstrate their understanding in different forms (verbally, pictorially, etc.)?

Additional Features: Does the task possess other noteworthy features?

A computer utility is being created that will prompt and keep track of judgments and comments and calculate summary scores for each criterion.

The procedure is intended to be usable with any set of learning goals that is adequately specific and clear. But it does not address two important issues: The general merit of a set of learning goals (beyond whether the goal statements lend themselves to this type of analysis) and the psychometric properties of a whole assessment instrument (beyond a summary of the types of tasks it includes).

In the Field

Experts in mathematics and science education have been helping Project 2061 to test and refine the analysis over the last two years. Numerous case studies will document how selected teams of educators are using the procedure to improve assessment tasks and create new ones so that they address standards. Some case studies will compare task scores with student interviews that explore what students understood the questions to be and what they meant by their responses. For example, one case study will analyze a task from the National Assessment of Educational Progress (NAEP) test in this way, a level of validation seldom done for large-scale tests. Another case study is being conducted on assessment tasks used in the Brownsville Urban Systemic Program in Texas for alignment with Texas Essential Knowledge and Skills (TEKS) standards. Educators will analyze the alignment of tasks from the 3rd grade Stanford 9 math test.

Project 2061 is actively seeking sets of science and mathematics assessment tasks for analysis. For more information contact assessment project manager Leah A. Bricker at lbricker@aaas.org or 202-326-7070. If you would like to learn more about aligning assessment to standards in a Project 2061 workshop, call 1-888-PDP-2061 or visit www.pdp2061.org.

Collaborating to Create Better Texts

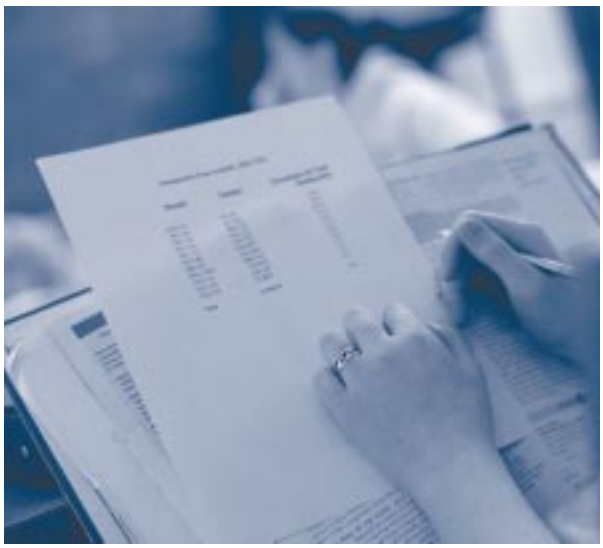
Project 2061 Hosts Conference on Science Textbooks

Many education stakeholders agree that the formula for teaching science to U.S. children must be rewritten to improve student performance. Science textbooks are an influential part of that formula. Poorly focused textbooks are a significant contributor to poor student performance, especially in the absence of good alternative materials. Therefore, teachers who must depend solely upon textbooks are unable to provide their students with a quality science education.

At a February 2001 conference sponsored by Project 2061, curriculum materials developers, education researchers, and commercial publishers convened to consider how to create a new generation of more effective science textbooks. The conference, "Developing Textbooks That Promote Science Literacy," came on the heels of rigorous evaluations by Project 2061 that found no middle-school science or high-school biology textbooks adequate in helping students meet national science education standards. The standards—Project 2061's *Benchmarks for Science Literacy* (1993) and the National Research Council's (NRC's) *National Science Education Standards* (1996)—identify specific ideas and skills that students should achieve at certain grade levels during their K–12 schooling. Project 2061 is providing the publishers of each book examined with a detailed report on why their publication received low marks.

The conference, sponsored by the National Science Foundation (NSF) and the David and Lucile Packard Foundation, brought together groups with distinct perspectives. The materials developers focus on creating innovative materials that take years to develop and just as long to learn to implement in the classroom. Commercial publishers, who must cater to their markets, maintain that school districts won't purchase books that are too difficult to use or that don't meet state adoption criteria, which may or may not align with national standards. And researchers are trying to uncover what they still don't know about how children learn.

The participants delved into the criteria and rationales used by Project 2061 in its textbook evaluations. The evaluations revealed fundamental problems such as disconnected facts, lavish illustrations that were needlessly complicated or inadequately explained, too much focus on technical terms that were easy to test but failed to communicate more important ideas, and failure to include or design activities that take account of commonly held student ideas. The conferees used CD-ROMs containing evaluations of curriculum materials on selected topics: matter and energy transformations and the kinetic molecular theory at the middle-school level and natural selection at the high-school level. They compared and contrasted sections of unsatisfactory textbook information with three satisfactory stand-alone science units developed by Michigan State University and the Michigan Department of Education. These three units were research-based and did not have the drawbacks of the textbooks.



Student Ideas Key to Learning

Project 2061's evaluations pointed to the importance of understanding and utilizing student ideas and misconceptions about science as learning opportunities. Jo Ellen Roseman, associate director at Project 2061, emphasized that incorporating student ideas is vital to successful learning. "One can ignore student[s'] ideas at their peril," Roseman warned. "Students have an amazing ability to retrofit and compartmentalize what goes on in school and leave relatively unaffected by the process. Our challenge is to find out what those ideas are and use them as learning tools."

continued on page 5

A Publisher's Perspective

Ellen Standafer, vice president of science product development for the publisher Holt, Rinehart and Winston, was a speaker and participant at Project 2061's conference "Developing Textbooks That Promote Science Literacy." Project 2061 staff member Cathy Tramontana recently interviewed Standafer via e-mail.

CT: What are the main challenges that publishers face when developing science textbooks in today's education environment?

ES: The development of materials requires that publishers serve an ever-growing number of special interests with different agendas. You can see one aspect of this problem in the diversity presented in various standards and frameworks documents for a single discipline or grade level. Even though state frameworks cite the use of *Benchmarks for Science Literacy* and the *National Science Education Standards* as a foundation, the spirit of those documents is often lost in the detail of the final product, because each framework reflects the group's current philosophy of science education.

Add the state adoption process, high-stakes testing, and the economic limitations of custom publishing to the problem, and the challenge for publishers becomes finding creative ways to integrate diverse and sometimes conflicting standards. In addition to state, national, and district frameworks, publishers have recommendations from discipline-specific professional organizations, such as the American Chemical Society, to contend with. Authors have their own criteria and the classroom teacher has his or her own concerns. Last, but certainly not least, the needs of students must play a big role.

Publishers are further challenged to bridge the gap between what teachers want for their classrooms and what their state or district mandates. Because teachers are placed in situations where they do not have the requisite backgrounds in the subjects they are teaching, some mandates seem impossible to follow. Budgetary limitations or a lack of facilities can make a mandate of 40% of instructional time for lab unrealistic.

CT: How does Holt, Rinehart and Winston handle these challenges?

ES: We look for as much commonality in standards as possible. We then decide if it makes sense to include the fringe topics or unusual program features. Obviously, these decisions are impacted by market size.

We also spend a lot of time and energy listening to classroom teachers talking about their schools, students, and instructional materials needs. Teachers turn to publishers to provide solutions for whatever instructional problems they face. For example, to support out-of-field teachers, the answer key teacher's edition is now a highly prescriptive set of daily lesson guides. To help science teachers deal with mandates to increase reading and math scores, we provide strategies to help them become reading and math instructors.

CT: After working with the Project 2061 instructional criteria at the textbook conference, in what ways do you find the criteria applicable to Holt's efforts to serve the needs of teachers and students? Are there any specific criteria that are especially pertinent to your work?

ES: Our focus since the conference has been to categorize the data on student misconceptions by product area. We have also begun drafting additional guidelines for artwork to prevent misconceptions related to our representations of phenomena. We can certainly be more accurate about how we use illustrations now that we can see how much the use of art can add to a student's learning.

CT: The second Project 2061 textbook conference in October 2001 will focus on the needs and perspectives of the textbook customer by eliciting the participation of state adoption committee members. What do you hope this second conference will address and accomplish with regard to state decision making about science textbooks?

ES: I hope the second conference creates an awareness that developing an array of unique curricula and standards is not the best way for states to solve the problems with science education in this country. We at Holt are looking at ways to better align with the AAAS criteria. Can the states do that as well? How can AAAS facilitate that goal?

We at Holt are looking at ways to better align with the AAAS criteria.

Better Texts *from page 3*

Several researchers and curriculum developers in attendance underscored the importance of attending to student ideas—at least on those topics where research exists. Norm Lederman, a science education researcher and teacher at Oregon State University and former high-school teacher, said one of the most common beliefs about the nature of science (or about how science works) held by many elementary- through high-school students—and by some of his own colleagues who teach science education—is that all science can be characterized by a single set and sequence of steps known as the scientific method. "Students believe that if you don't follow this precise method, you are not doing science," said Lederman. Jim Minstrell, currently a researcher at Talaria, Inc., who taught high-school mathematics and science for more than thirty years, discussed a program he has developed called Diagnoser, which helps teachers identify and qualify what he calls "commonly held ideas" or "facets" about topics in physical science. And Kathleen Roth, a curriculum materials developer and researcher at Michigan State University, described the key role that eliciting and guiding student thinking about scientific concepts played in the development of her research-based life science unit, *Food for Plants*.

National Standards versus State Adoption Criteria

In addition to learning about the work of researchers and curriculum materials developers, conference participants gained a greater appreciation for the challenges faced by publishers, who must grapple with the mission to create exceptional student textbooks and teacher guides based on the national standards while meeting state adoption criteria and satisfying the wants and needs of the teachers. Publishers, while striving to develop materials that best serve the educational needs of the students, say they are constrained by what teachers and schools will buy. "All of us are concerned about the paradox of standards versus adoption," said Rodger Bybee, executive director of Biological Sciences Curriculum Study, developers of secondary-school and college-level programs in the life sciences. Bybee, a key player in the development of the NRC's *National Science Education Standards*, pointed out that many state and local school districts have adoption committees that dilute the standards by adding and subtracting content without considering the consequences. These rewritten standards become the basis for textbook adoption criteria, and publishers must meet these criteria if their books are to be included on the list of materials that schools can purchase. Thus, as Robert Todd of Holt, Rinehart and Winston stressed, teachers need to buy into the national standards criteria in order to drive the market for textbook changes.

Changing the Textbook Market

Lou-Ellen Finn, professional development coordinator for Northwestern University's Center for Learning Technologies in Urban Schools, observed that future conferences should create more opportunity for a dialogue among publishers, researchers, and curriculum materials developers. "We've got publishers over here with their concerns, researchers over there with their concerns, and unless we can leverage everybody's expertise and get us all on the same page working toward the same thing—student learning—it's never going to happen," said Finn.

Some of the conferees agreed that collaborations between researchers, curriculum materials developers, and publishers could help create a market among teachers and policymakers for standards-based curriculum and textbooks. George Nelson, director of Project 2061, pointed out that the education community is steeped in traditions and some of these need to be revisited in order to improve student learning. "One of our long-term goals at Project 2061 is to change the market and start new traditions."

A second conference will be held in October 2001 to continue working toward science textbooks that promote science literacy. Among the participants will be key decision makers who develop and uphold state adoption criteria.

Project 2061

Products and Services

Available from Oxford University Press

The following items can be ordered by mail or telephone from:
Oxford University Press, Ordering Department
2001 Evans Road, Cary, NC 27513
Telephone: 1-800-451-7556
Oxford University Press U.S. web site: http://www.oup-usa.org

Science for All Americans

ISBN 0195067711, 1989

Project 2061's landmark publication outlining what high-school graduates should know and be able to do in science, mathematics, and technology. $14.95 (also available in Spanish and Chinese)

Benchmarks for Science Literacy

ISBN 0195089863, 1993

This companion to *Science for All Americans* presents a set of science, mathematics, and technology learning goals for K–12 students. $26.50 (also available in Spanish and Chinese)

Resources for Science Literacy: Professional Development

ISBN 0195108736, 1997

This book and CD-ROM combination offers science educators a better understanding of what science literacy is, what it requires of students, and how teachers can help students achieve it. $49.95

Blueprints for Reform: Science, Mathematics, and Technology Education

ISBN 0195124278, 1998

Blueprints for Reform provides a starting point for reforming the education system, outlining changes needed in 12 specific areas to improve learning in science, mathematics, and technology. $17.95 (also available in Chinese)

Designs for Science Literacy NEW

ISBN 0195132785, 2001

This new guide provides techniques and strategies for designing a K–12 curriculum that is properly aligned with established learning goals, logically connects all subjects and grade levels, and reduces the number of topics covered. A companion CD-ROM demonstrates the functions that curriculum design software can perform and provides aids to assist in carrying out recommendations from the book. $49.95

Available from AAAS/Project 2061

The following items can be ordered by mail or telephone from:
AAAS/Project 2061 Ordering Department
1333 H Street, NW, 8th Floor, Washington, DC 20005 ◀ NEW ORDERING ADDRESS!
Telephone: 1-888-737-2061 Fax: 202-842-5196

UPDATE 2001–2002

This **FREE** publication describes Project 2061's latest research and programs.

Dialogue on Early Childhood Science, Mathematics, and Technology Education

ISBN 0871686295, 1999

This collection of essays by early childhood education experts discusses the latest findings on teaching science, mathematics, and technology to preschool children. $12.95

Middle Grades Mathematics Textbooks: A Benchmarks-Based Evaluation

ISBN 087168635X, 2000

Features the results of Project 2061's analysis of middle grades mathematics texts and a detailed description of the procedure used for the analysis. Includes a CD-ROM that allows side-by-side comparisons of data. $89.00

Atlas of Science Literacy NEW

ISBN 0871686686, 2001

A collection of "strand" maps for nearly 50 topics that are important for literacy in science, mathematics, and technology provides new perspectives for educators. Each map depicts how students might move from one level of knowledge and understanding to the next, shows connections among major ideas, and suggests the sequence in which the ideas should be learned. $49.95

Project 2061 Professional Development Programs

Project 2061 creates customized programs that clarify the intent of national, state, or district standards and help educators align curriculum and assessment to those standards. Workshops that train educators to evaluate and select textbooks and other curriculum materials are also available. For more information, call 1-888-PDP-2061 or visit www.pdp2061.org.

Director's Notes

Comfortable on the Edge

Think of what can be understood about the world as a small, slowly expanding island in an infinite sea of mystery. Scientists, mathematicians, engineers, and scholars and visionaries of all kinds are drawn to the edge between what is understood and what has not been observed or explained. It's where they are comfortable and where they enjoy "the pleasure of finding things out," as the Nobel prize-winning physicist Richard Feynman described it.

Education is an exciting field in part because we know so little. Our understanding of how people learn mathematics, science, and technology is still rudimentary but already opens up some wonderfully complex questions. Finding answers to these will keep researchers busy for decades to come. In the meantime, like our colleagues in science and engineering, we can use the knowledge we have today to do what we can, fully expecting that things will change tomorrow.

Research as a Guide

Lately we've been impressed with *How People Learn: Brain, Mind, Experience, and School*, a new book from the National Research Council. It provides some examples of how we might put knowledge about learning into practice and build on that knowledge with a carefully planned research agenda. For instance, while benchmarks and standards define the core content and describe the outcomes of good instruction, they do not specify the instruction itself. Without research as a guide, the education community is easily swept along by the next wave of "innovation." Too often educators simply adopt the latest instructional techniques—"technology-based" or "inquiry-based hands-on minds-on problem-based peer-led open-ended learning cycles"—rather than take the time to understand or engage in the research that may, or may not, support those techniques in a given application. *How People Learn* provides a solid research foundation for incorporating knowledge of learning into the instructional approaches that are appropriate to the learning goals, the learners, and the situation.

Project 2061's approach to analyzing curriculum materials offers another example of how research findings can be operationalized. Drawing extensively on what was currently known about student learning, we developed a set of criteria to judge the quality of a textbook's instructional framework. Since then, other research teams have been using the criteria in a curriculum materials design setting to see if this new application can work. We're taking a similar approach to assessment, again turning to the research base to help define what it means for an assessment task to align to standards and to think about the characteristics of effective assessment. With both efforts, we make the best use of today's research and build on it to expand our island of knowledge.

What Don't We Know?

There are so many open questions. Do adults learn differently from children? What kinds of professional development experiences help teachers to understand and implement new standards-based materials? What kinds of pre-service experiences prepare teachers to use research on learning to best advantage in their classrooms? What impact do these kinds of changes have on student learning? How can curriculum materials be developed to serve diverse populations of students? Which of our own criteria for judging instructional quality are most important? Are some expendable and are new ones needed? The list could go on and on.

These are hard questions both technically and intellectually. They take us beyond our usual "reform" activities of re-rewriting our standards and uncritically sampling from the smorgasbord of new instructional tricks while trying to teach the same thin content. They force us to the boundaries of our knowledge and challenge today's ideas. But for us at Project 2061—comfortable on the edge—that's the fun.

George D. Nelson
Director

Our understanding of how people learn mathematics, science, and technology is still rudimentary but already opens up some wonderfully complex questions.

FYI

Introducing…

Project 2061 welcomes three new staff members. **Leah Bricker** has joined the project as the senior program associate responsible for coordinating the project's assessment task analysis. Bricker previously served as the science curriculum program coordinator for the Indiana Department of Education. A former middle-school science teacher, Bricker holds an M.S. in Interdisciplinary Biology from Indiana University-Purdue University. Having completed a year of Ph.D. course work in Science Education at Purdue University, she plans to continue her studies at the University of Maryland in the fall. Joining Project 2061 as an administrative coordinator is **Michele Douglas**, who has more than 13 years of experience in administration, project coordination, and customer service. Her previous employers include Caliber Associates and the World Bank. Technology specialist **Ed Krafsur** is currently developing a web-based interactive utility to accompany the assessment task analysis procedure. His previous experience includes working in information technology support at the Iowa State University College of Agriculture.

Paul Kimmelman Joins PDP

Dr. Paul Kimmelman has joined the staff of Project 2061 Professional Development Programs (PDP) as a senior consultant. In this capacity, he represents the project at public events and speaking engagements, helps school districts design research-based professional development programs, and assists PDP with strategic planning. Dr. Kimmelman brings to Project 2061 a broad perspective on science and mathematics education reform. He recently retired after a long career as an educator that included 15 and a half years as a superintendent. He served as president of the First in the World Consortium in Chicago, IL, which was the first group of U.S. school districts to participate in the Third International Mathematics and Science Study (TIMSS). He also served as a member of the National Commission on Mathematics and Science Teaching for the 21st Century (the "Glenn Commission") and is regularly called upon to testify to the U.S. Congress on educational policy.

Evaluation of Algebra Textbooks Now On-Line

The project's evaluation of 12 widely used algebra textbook series is now available on the Project 2061 web site at www.project2061.org/tools/textbook/algebra. *Algebra Textbooks: A Standards-Based Evaluation* allows users to (1) browse in-depth evaluation reports on each of the 12 textbook series; (2) view two different data sets at the same time, either from the same book or from two different books; (3) examine the Project 2061 analysis procedure for mathematics curriculum materials; (4) read an explanation of the evaluation's methods; and (5) review example scenarios for applying the results of the evaluation. Teachers and adoption committees can use the evaluation to make initial decisions about textbooks or to identify the strengths and weaknesses of textbooks already in use.

Organizations Applaud New Science Standards in Kansas

The AAAS, the National Academy of Sciences and National Research Council, and the National Science Teachers Association have released a joint statement of support for the Kansas State Board of Education's decision to adopt new K–12 science standards. The February 2001 decision overturned the actions of the previous Board by reinstating the study of the origins of life and the cosmos to the Kansas state science standards. As part of their statement, the three organizations have granted copyright permission to reference or use text from their documents in the new standards, permission the groups had denied the previous Board. To read the joint statement, see www.nsta.org/pressrel/kansas_statement.asp.

2061 *today*

American Association for the Advancement of Science
1333 H Street, NW
PO Box 34446
Washington, DC 20005

Return Service Requested

Nonprofit Organization
U.S. Postage
PAID
Washington, DC
Permit No. 5676