# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>ix</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xi</td>
</tr>
<tr>
<td>FOREWORD</td>
<td>xiii</td>
</tr>
<tr>
<td>Michael Padilla</td>
<td></td>
</tr>
<tr>
<td>ABOUT THE EDITORS</td>
<td>xvii</td>
</tr>
<tr>
<td>PART I: PISA SCIENCE 2006</td>
<td></td>
</tr>
<tr>
<td>Chapter 1</td>
<td>3</td>
</tr>
<tr>
<td>PISA: An Introduction and Overview</td>
<td></td>
</tr>
<tr>
<td>Ross Turner</td>
<td></td>
</tr>
<tr>
<td>Chapter 2</td>
<td>15</td>
</tr>
<tr>
<td>PISA 2006: An Assessment Framework for Scientific Literacy</td>
<td></td>
</tr>
<tr>
<td>Rodger W. Bybee and PISA 2006 Science Expert Group</td>
<td></td>
</tr>
<tr>
<td>Chapter 3</td>
<td>27</td>
</tr>
<tr>
<td>PISA 2006 Test Development and Design</td>
<td></td>
</tr>
<tr>
<td>Barry J. McCrae</td>
<td></td>
</tr>
<tr>
<td>Chapter 4</td>
<td>39</td>
</tr>
<tr>
<td>PISA: Frequently Answered Criticisms</td>
<td></td>
</tr>
<tr>
<td>Raymond J. Adams</td>
<td></td>
</tr>
<tr>
<td>PART II: INTERNATIONAL COMPARISONS</td>
<td></td>
</tr>
<tr>
<td>Chapter 5</td>
<td>49</td>
</tr>
<tr>
<td>PISA Science 2006: International Results</td>
<td></td>
</tr>
<tr>
<td>John Charles Cresswell</td>
<td></td>
</tr>
<tr>
<td>Chapter 6</td>
<td>59</td>
</tr>
<tr>
<td>Knowledge Of and About Science</td>
<td></td>
</tr>
<tr>
<td>Rolf V. Olsen and Svein Lie</td>
<td></td>
</tr>
<tr>
<td>Chapter 7</td>
<td>69</td>
</tr>
<tr>
<td>What Scientific Knowledge Remains When the Rest Is Forgotten?</td>
<td></td>
</tr>
<tr>
<td>Ewa Bartnik and Barbara Ostrowska</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 8
What Lies Behind Finnish Students’ Success in PISA Science? 79
Merike Kesler and Jari Lavonen

Chapter 9
Overcoming Challenges and Succeeding in PISA Science 2006 91
Robert Laurie

Chapter 10
Everyday Conceptions Across the World 101
Svein Lie and Rolf V. Olsen

Chapter 11
A Perspective on U.S. Science Teaching and Learning 111
Manfred Prenzel and Tina Seidel

Chapter 12
Improving Science Teaching and Learning 117
Paulina Korsnáková, Barry J. McCrae, and Rodger W. Bybee

Chapter 13
Windows Into High-Achieving Science Classrooms 123
Joseph A. Taylor, Molly A. M. Stuhlsatz, and Rodger W. Bybee

Chapter 14
The Importance of Aligning Teaching and Assessment 133
Andrée Tiberghien

Chapter 15
PISA 2006 Assessment of Attitudes Toward Science 139
Yasushi Ogura

Chapter 16
What Science Do Students Want to Learn? 149
Barry J. McCrae

Chapter 17
Teaching and Learning Science: PISA and the TIMSS Video Study 163
Sue Thomson

PART III: IMPLICATIONS FOR SCIENCE TEACHING

Chapter 18
Seeing the U.S. Education System Through the Prism of PISA 179
Andreas Schleicher
Chapter 19
Teaching Science to Achieve Scientific Literacy 187
Peter J. Fensham

Chapter 20
Designing a Science Curriculum to Enhance Students’ Scientific Literacy 203
Robin Millar

Chapter 21
Assessing PISA 2006 Scientific Competencies 215
Marietta Bloch and Robert Laurie

PART IV: PISA AND SCIENCE EDUCATION
Chapter 22
Scientific Literacy: Implications of PISA Science 2006 for Teachers and Teaching 227
Rodger W. Bybee and Barry J. McCrae

APPENDIX
PISA 2006 Science Units 249

INDEX 281
As science teachers think about preparing students for today’s world, they may well ask, “What should I teach that will enable my students to fully participate in a society in which science and technology play a significant role?” Knowledge about how students respond to information, problems, and issues that require some understanding of science and technology would help science teachers answer this question and make changes in their courses and teaching.

The Programme for International Student Assessment (PISA) provides such information about 15-year-olds that can be both insightful and useful for science teachers. PISA is a collaborative effort undertaken by member countries (and many nonmember countries) of the Organisation for Economic Co-operation and Development (OECD), which has its headquarters in Paris, France.

PISA measures 15-year-olds’ capabilities in reading literacy, mathematical literacy, and scientific literacy every three years. PISA was first implemented in 2000, and the most recent results are for the 2006 assessment. Each survey assesses one subject area in depth. The other two subjects also are assessed at the same time but not in the same breadth and depth as the primary or main domain. In 2000 reading literacy was the main domain, in 2003 it was mathematics, and in 2006 it was science. PISA 2003 also assessed cross-curricula problem solving.

PISA uses the term literacy within each domain or subject area to indicate a focus on the application of knowledge and abilities. Literacy refers to a continuum of knowledge and abilities; it is not a designation of a condition that one has or does not have. PISA assessments do not provide data to determine who is literate or illiterate.

The assessment measures how well 15-year-old students are prepared to meet the challenges they may encounter in life. At age 15, students in most OECD countries are approaching the end of compulsory schooling, so the assessment apprises professional educators, policy makers, and parents of the knowledge, skills, and attitudes developed during 10 or so years of education.

Rather than the curricular orientation of the Trends in International Mathematics and Science Study (TIMSS), PISA provides a unique and complementary perspective by focusing on the application of knowledge in reading, mathematics, and science to problems and issues posed in real-life contexts. The results of PISA answer the question “Considering schooling and other factors, what knowledge and skills do students have, and can apply, at age 15?” The achievement scores from PISA represent a “yield” of learning at age 15, rather than a measure of the attained or learned curriculum at grades 4 or 8, as is the case with TIMSS.

The chapters in this book use the PISA 2006 scientific literacy framework and results from the survey as their central focus. The 2006 assessment represented the first time that science was the primary domain in PISA, with the competencies of
scientific literacy as the major focus. PISA Science 2006 has significant implications for science teachers and teaching. In order to respond to demands of the 21st century, including implications of a global economy and environmental concerns, science teachers can use PISA results to inform their decisions as they answer questions about what to teach and why those understandings, skills, and attitudes they are teaching will contribute to a student’s scientific literacy.

The Book’s Organization
After a foreword by Michael Padilla, NSTA president 2005–2006, the chapters are grouped in four sections. Part I introduces PISA Science 2006. The first chapter provides an overview of PISA, and subsequent chapters present the framework that defines and clarifies scientific literacy, describe the test’s design and development, and discuss frequently answered criticisms. Our intention for this section is to present a detailed introduction to the PISA survey for 2006. All the authors were directly involved with the design, development, implementation, scoring, and reporting of the results.

Part II begins with an overview of the international results and then presents a series of short essays addressing important themes for science teachers. The authors of these chapters come from a wide range of countries and were involved in various capacities with PISA Science 2006. Many of them were members of the Science Expert Group that oversaw the assessment. The authors bring their unique perspectives to the essays and provide science teachers with insights and international comparisons on topics such as the importance of knowledge about science, why high-achieving countries (e.g., Finland and Canada) are so successful, students’ misconceptions, improving science teaching and learning, students’ attitudes and interests, and a comparison of PISA 2006 outcomes with the results from the TIMSS 1999 Video Study of eighth-grade science teaching.

Each chapter in Parts I and II begins with a quotation from the PISA 2006 science framework that establishes the theme for the chapter. For Part III we asked international experts to describe the implications of PISA 2006 for essential components of science teaching: goals, curriculum and instruction, and assessment. These chapters also have more detail about topics of specific importance for science teachers. All authors have expertise in science education and extensive experience with their topics.

The book concludes with a chapter on the relationship between PISA and science education policies, programs, and practices. In this chapter we make a statement about scientific literacy and take a larger perspective on science education. Citing the chapters in the book, we present a case for scientific literacy and the need for coherent educational policies, school programs, and classroom practices if we wish to attain higher levels of scientific literacy and a workforce prepared for the 21st century.

Each chapter is relatively self-contained, so the chapters can be read independently of each other except, perhaps, for the occasional need to consult Part I for some PISA background information. This means that there is some repetition between chapters, especially in references to framework material, but we deemed this unavoidable. Occasional cross-references to other chapters of the book point the interested reader to further, related information.

Rodger W. Bybee and Barry J. McCrae, Editors
The ideas and insights expressed in *PISA Science 2006: Implications for Science Teachers and Teaching* were developed by the various authors during several years as work progressed on the PISA 2006 survey. We thank each of them for accepting our invitation to contribute to this publication and giving freely of their time in doing so.

The Programme for International Student Assessment is a large and complex undertaking. Science is only one component of the survey. PISA 2006, like PISA 2000 and 2003, was implemented internationally on behalf of the OECD by an international consortium led by the Australian Council for Educational Research (ACER). We thank ACER for the many ways in which they have supported the preparation of this book, including the provision of writing time and travel assistance to the editors through its Professional Writing Program fund.

Ray Adams of ACER and the University of Melbourne has been the international project director since PISA’s inception in 1998. His commitment to the dissemination of PISA findings provided the original impetus for this publication. We acknowledge his outstanding leadership of PISA and thank him for his ongoing encouragement and support for our work.

Strong support for the work also came from the Organisation for Economic Co-operation and Development (OECD), in particular from Andreas Schleicher (head of the Indicators and Analysis Division, Directorate of Education) and John Cresswell. Reports prepared by or on behalf of the OECD provided the foundation for much of the content of this book.

We also recognize those groups that contributed to preparation of the PISA 2006 science test material. Here we acknowledge ACER, Westat and the Educational Testing Service (ETS) in the United States, Citogroep in the Netherlands, the Institute of Teacher Education and Learning at the University of Oslo in Norway, Leibniz Institute for Science Education at the University of Kiel in Germany, and the National Institute for Educational Policy Research in Japan. We also thank Ross Turner, Ron Martin, and Maurice Walker of ACER for their thoughtful critiques and recommendations that improved specific chapters and the book in general, and Alla Berezner and Eveline Gebhardt of ACER for their support in performing secondary analyses of PISA data for various chapters.

We had wonderful support throughout the development of this book. Byllee Simon at the Biological Sciences Curriculum Study (BSCS) provided administrative assistance in the early stages of the project. Patricia Bybee helped with the organization of work and coordination between authors in 11 countries, the editors in the United States and Australia, and the National Science Teachers Association (NSTA). At NSTA, we are thankful for the support and understanding of Claire Reinburg and Betty Smith.

*Rodger W. Bybee and Barry J. McCrae,* Editors
Why Is International So Important and What Role Do International Comparisons Play?

As I write this foreword in April 2008, I am looking at today’s edition of the New York Times. Of 6 articles on page one, 4 relate to international and only 2 to domestic issues. These include articles on trade with Columbia, Asian inflation, the Mexican border fence, and stability in Iraq. Similarly, 5 of 12 articles in the news section of the USA Today website focus directly on international topics. More importantly, even the obviously domestically oriented articles (for example, the presidential election and fuel prices for airliners) have international implications. Clearly, this one-day snapshot communicates the trend that global issues drive our world more strongly today than ever before.

Issues related to economic competitiveness, immigration, and relations with other countries are commonplace headlines. We read about factories closing in the Midwest and moving to Bangladesh, about avian flu, and about poor U.S. performance on international tests of mathematics and science. However, the public seldom sees the connection between factory closings and education. Still it is just this broader picture that is so important to truly understanding the significance of international comparisons of student achievement. Let me outline a few central themes found in the news that bolster the case for why an international focus on science education is so critical to the global business, politics, and the future. I will then attempt to link these broader societal themes back to international comparisons of student achievement and to the Programme for International Student Assessment (PISA) Science 2006 results.

The International Picture

Economic issues dominate national and international politics. Most nations promote economic well-being as a central goal for the future and cite the importance of education (and specifically science, math, and technology education) in achieving this goal. In The Elephant and the Dragon: The Rise of India and China and What It Means for All of Us, Robyn Meredith (2007) expands the thesis, first popularized by Tom Friedman in The World Is Flat (2005), that an educated workforce is the basis of future economic competitiveness and prosperity in both developed and developing nations. Both Meredith and Friedman argue that, as a result of technology, business is interconnected, is hypercompetitive, and requires international cooperation. The winners will be those societies that have the most well-educated citizens. While the news media write about how the United States is vying with China, Vietnam, and Sri Lanka for manufacturing jobs, competition for education-intensive, white-collar jobs with India is more significant. Robyn Meredith states that there always will be a
country with wages lower than those in the United States, and our future is dependent on educating individuals for higher-paying jobs that require more education.

A second theme widespread in the news media relates to immigration. While building a fence on the Mexican border is prominent in North American media, the same core issue of workers migrating illegally to wealthy nations is almost universal. I have led several study abroad trips for teachers to Central America. One of our core activities on these trips is to study the immigration problem in Costa Rica. Poor Nicaraguans provide a much-needed labor force in Costa Rica, but they also place burdens on the Costa Rican education and social services infrastructure. Talk is often harsh about how these illegal immigrants burden the system, refuse to assimilate, and should be deported immediately. Sound familiar? While we sometimes think the U.S. situation with Central Americans is unique, similar scenarios are playing out across Europe and in numerous other countries on almost all continents.

Although immigration issues seem unrelated to science and mathematics, let’s not forget that U.S. law requires states to educate children regardless of their citizenship. Moreover, children born of illegal immigrants in this country are legal citizens of the United States. If we want these children to be productive citizens who contribute to our political and economic well-being, then we must educate them, especially in mathematics and science. This will necessitate dramatic changes in the teaching of all subjects, including mathematics and science, and will demand a significantly improved teaching force.

Resolution of global problems is yet a third theme in the news—one that is often related at its core to science, mathematics, and technology. Twenty-first century global issues such as AIDS, global climate change, clean air and water, and the provision of adequate and safe food supplies are not only related to science but also affect every nation on the planet. It is clear that resolution of these problems will take both a public understanding of the issues and the technological know-how of scientists and mathematicians.

Collaboration and cooperation is the fourth and last theme, with almost all nations agreeing that these processes will be a big part of solutions to the first three global problems. In fact, this last theme is the one that ties all the others together. Both collaboration and cooperation are necessary for individuals and nations to compete in tomorrow’s workforce, and both will be required to provide resolution of immigration and global environmental issues.

Vivien Stewart (2007) of the Asia Society summarized the importance of the international themes well by describing how they will affect tomorrow’s workforce. She stated that U.S. high school graduates of the future will

- sell to the world;
- buy from the world;
- work for international companies;
- manage employees from other cultures and countries;
- collaborate with people all over the world in joint ventures;
compete with people on the other side of the world for jobs and markets; and

- tackle global problems, such as AIDS, avian flu, pollution, and disaster recovery.

Regardless of whether you think of the future through the filter of the economy and tomorrow’s workforce or whether you think of the political consequences of poor international relations, education is clearly the keystone to improving the world. Without an increasingly educated populace, there is little hope of a better world. So how can we connect this broad idea to international assessment in mathematics and science?

Relating International Themes to International Comparisons
At their worst, international educational comparisons are a horse race for which results are reported by the press and not much else. They cause the requisite wringing of hands among the public and editorial outcry in the newspapers and on television. Teachers and school systems often feel demoralized and despondent as a result, but go about their jobs in spite of it.

At their best, international comparisons like TIMSS (Trends in International Mathematics and Science Study) and PISA allow nations to learn how well (or poorly) they are doing relative to peer nations and provide a basis for benchmarking improvement. They stir educational policy makers and politicians alike to examine policies and practices. They motivate professionals to find out what makes certain countries successful.

If you travel abroad much, you know most countries take these outcomes very seriously. They know where they stand and have begun to figure out how to improve. I am continually surprised when visiting nations like Japan, Korea, and Singapore to learn what changes they are making to their educational systems. These top-achieving countries are not resting on their laurels but rather finding ways they can improve. They are looking for new ways to motivate students and teachers. They know that being top scoring in any achievement test is not the sole outcome of schooling and are trying to improve other outcomes as well. They are investigating what other countries are teaching, how they are teaching it, and how it can be adapted to their needs. They are embarking upon long-term professional development plans to improve their country’s teaching and learning. In fact, no nation that I have visited is satisfied with its achievement, and all know that the stakes are high. Even those countries that have not participated in either TIMSS or PISA are, nonetheless, knowledgeable about and aware of the results.

For years, the educational system within a specific country acted alone and separately from that of other countries. Each responded to pressures and issues within its own political systems, but not much else. So the changes we are seeing are quite remarkable and very positive. For the first time, countries have data that can be used by educators to leverage improvements in their system. And they are using this knowledge to influence politicians. Moreover, many of the politicians themselves are aware of how education connects to the political and economic future, and many of the visionary ones are pushing the systems themselves.
Let’s look a little more closely at TIMSS and PISA and what they offer. While both provide important data for accomplishing educational improvement, PISA results, as you will learn throughout this book, provide a breadth of assessment that goes beyond basic content knowledge. The PISA literature states that it is a measure of literacy within the three subjects it assesses—science, mathematics and reading—and focuses on “real-life problems and lifelong learning.” PISA items address “what students can do with what they learn at school and not merely whether they can reproduce what they have learned.” TIMSS in contrast is a measure primarily of what students know and as such is a more traditional curriculum-based assessment of school subjects. Although both have raised the awareness and importance of science and mathematics throughout the world, they have done so in different ways.

More importantly, both TIMSS and PISA have given international comparative educational research a new legitimacy and validity among the public. Politicians and educated citizens now take these results seriously and are placing demands on schools never before made. I think these demands, while difficult in the short term, are necessary if we are to see all nations reach a level of development only attained by a few today.

In summary, the world is changing and will continue to do so. Yet the themes described in this foreword are likely to remain the same. Economic development, immigration, and global issues related to science, mathematics, and technology are likely to persist and even increase. Collaboration and cooperation are likely to be more meaningful and necessary to future generations. The only way of dealing with these issues will be through improved education, especially in science and mathematics, with all players in the arena doing their part. In response to this reality, the National Science Teachers Association convened an international task force that proposed a dramatic increase in the association’s international focus and outreach—including international conferences, study abroad trips for teachers, and collaboration with science teacher organizations worldwide (NSTA 2005). These new policies were adopted in 2005 and are quickly being implemented. The role that international comparisons of achievement play in the overall process is also significant. With this context in mind, I am confident that leaders will find the chapters of this book to be required reading.

Michael Padilla
NSTA President 2005–2006
Clemson University

References
About the Editors

Rodger W. Bybee was chair of the PISA 2006 Science Expert Group. He is a former executive director of the Biological Sciences Curriculum Study (BSCS) and of the Center for Science, Mathematics, and Engineering Education at the National Academies of Science, Washington, D.C.

Barry J. McCrae is a principal research fellow and leader of the Mathematics and Science test development team at the Australian Council for Educational Research. Before joining ACER in 2001, he was involved for many years with the preservice and post-service training of mathematics and science teachers at the University of Melbourne, where he now holds an honorary appointment of principal fellow. He managed framework and test development for the PISA 2006 assessment of scientific literacy.
PART I

PISA Science 2006
1 PISA: An Introduction and Overview

Ross Turner

Ross Turner is a principal research fellow in the National and International Surveys research program of the Australian Council for Educational Research (ACER). Building on earlier roles in secondary mathematics teaching, mathematics curriculum development, student assessment, and the monitoring, evaluation, and reporting of assessment outcomes, he has coordinated the development and implementation of the mathematics component of the international PISA project since 2000. He now also fills a broader role managing ACER’s contribution to the international consortium that implements PISA on behalf of the Organisation for Economic Co-operation and Development (OECD).

Introduction

For most science teachers, the term PISA evokes images of a leaning tower in an Italian city, but PISA also is an acronym for a major international assessment of students. In 2006 the primary assessment domain of PISA was science. This chapter provides a general introduction to PISA 2006.

What PISA Is and Does

The Programme for International Student Assessment (PISA) is an international comparative educational survey carried out for the Paris-based Organisation for Economic Co-operation and Development (OECD). Established in the late 1990s, PISA periodically surveys the competencies in mathematics, science, and reading of 15-year-old students in OECD member countries and in an increasing number of nonmember countries. The students are approaching the end of compulsory schooling in most participating countries, and school enrollment at this level is close to universal in almost all OECD countries.

PISA surveys student competencies every three years. The first survey took place in 2000, and the second in 2003. The international results of these surveys have been published through the study’s initial international reports (OECD 2001, 2003a; OECD 2004a, 2004b). A number of other reports addressing specific aspects of PISA have been published by the OECD and can be accessed through the OECD’s website at www.pisa.oecd.org.

The most recent PISA survey, which emphasized science, was conducted during 2006. Results were published in December 2007 (OECD 2007a, 2007b). For each survey, mathematics, science, or reading is chosen as the major assessment domain.
The other two areas are assessed as minor domains. In 2000 the major domain was reading, and in 2003 it was mathematics.

**Student Tests**

The PISA assessments center on literacy. This perspective focuses on the extent to which students can use the knowledge and skills they have learned and practiced at school when confronted with life situations and challenges for which that knowledge may be relevant. PISA seeks to assess the extent to which students can use their reading skills to understand and interpret various kinds of written material they are likely to meet in their daily lives; the extent to which students can use their mathematical knowledge and skills to solve various kinds of mathematics-related challenges and problems; and the extent to which students can use their scientific knowledge and skills to understand, interpret, and resolve various kinds of scientific situations and challenges.

In addition to these three central areas that are studied in each PISA survey, the survey from time to time allows for the assessment of additional cross-curricular competencies as participating countries see fit. For example, in the 2003 survey an assessment of general problem-solving competencies was included. In 2006, PISA included an optional computer-based assessment of scientific literacy, though this option was fully implemented only by three countries (Denmark, Iceland, and Korea).

**Student and School Questionnaires**

In each survey, students complete a questionnaire designed to gather relevant background data about their personal characteristics, opinions, preferences, and aspirations; some characteristics of their home and family environments; and selected features of their school environments. School staff complete a short questionnaire about various aspects of organizational and educational provisions in schools. This background and contextual information is collected to generate a detailed study of factors within and between countries that are associated with varying levels of reading, mathematical, and scientific literacy among the 15-year-old students of each country.

**How PISA Is Organized**

The organizational structure of PISA is shown in Figure 1.1. The PISA project is overseen by a secretariat within the OECD in Paris. Policies that guide the project are set by the PISA Governing Board (PGB), an OECD committee comprising delegates and observers who are largely senior educational administrators from the participating countries. The PGB meets twice each year.

Each participating country has a national project implementation center. A national project manager (NPM) is appointed to coordinate all activities at the national level. Typically, the NPM works closely with the country’s PGB member to establish a national perspective on policy matters, on matters related to the implementation of PISA, and on the analysis and reporting of outcomes that may be of particular
relevant to the country. Generally the national center has a small team of staff working on project development, implementation, and reporting at the national level.

The project is implemented internationally by a contractor appointed by the OECD. An international consortium led by the Australian Council for Educational Research (ACER) was the sole contractor for each of the first three PISA survey cycles and is one of two contractors for PISA 2009. ACER’s consortium partners for PISA 2006 were the National Institute for Educational Measurement (Citogroep) in the Netherlands, Westat and the Educational Testing Service in the United States, and the National Institute for Educational Policy Research in Japan. Ray Adams of ACER and the University of Melbourne has been the international project director since PISA commenced.

To carry out the range of project implementation tasks required, ACER has marshaled the services of teams of people with expertise in a variety of areas. Some of this expertise resides in the staff available through ACER’s consortium partner organizations. In addition, the consortium works with several key groups and individuals.
that include a number of subject expert groups comprising internationally recognized experts in the specific assessment domains and technical specialists in such areas as sampling, test and questionnaire item development, translation, and statistical analysis.

Most important, consortium staff work very closely with the personnel involved at the national level. The connection between national project managers and the international consortium is critical to the success of PISA. Each depends on the other to ensure successful implementation of the project. National centers provide the consortium with information about conditions and constraints that exist in the country, with feedback regarding the various project elements that are being developed, and with important advice on how the project can best be implemented in that country. The consortium provides national centers with information about project implementation standards and requirements, draft materials for national consideration and feedback, and training and materials to facilitate project implementation.

Other consultative mechanisms have been used, for example, the Science Forum for PISA 2006. At this open forum participating countries could nominate national experts to directly represent the interests and views of the country in considering certain detailed technical aspects of the project. The Science Forum allowed for a wider base of expert input than was possible through the PISA 2006 Science Expert Group. The Science Forum considered priorities and issues at the time the science framework was being conceptualized, and it provided important input in the development of survey material related to the assessment of science and the assessment of student attitudes to science.

Participating Countries
With the completion of PISA 2009, 65 countries will have participated in one or more PISA surveys. They are listed in Table 1.1.

Technical Aspects of PISA
PISA employs rigorous scientific survey procedures that fall into several different technical areas. These are briefly introduced here. Some will be discussed in greater detail later in this book.

Test Development
PISA assessments are based on frameworks (OECD 2000, 2003b, 2006) that are negotiated and agreed on among participating countries. The frameworks drive what is to be assessed—they define the constructs that underpin the assessments in each of the different domains that are tested, and they give direction to the test item developers. ACER has led test development for the first four PISA surveys and has drawn on test development expertise from other organizations. Participating countries also contribute test items and ideas for items, contributions that are further developed and supplemented by the consortium’s expert test developers, using state-of-the-art cognitive laboratory procedures. In this way a large number of test items are developed for possible use in the PISA survey. They are presented in...
a variety of test item formats, including multiple-choice and constructed-response formats. Routitsky and Turner (2003) show the importance of using a mixture of item formats in PISA.

In the year before the main survey, an extensive field trial takes place in which all test items proposed for selection for the main survey are administered to students in each participating country. The data from the field trial, together with feedback received from countries on the proposed items (for example, an item’s content and its structure, and potential translation or cultural issues), are used as the basis for selecting the final main survey test item sets. As a result, the final item set provides the best possible coverage of the relevant framework and has been shown empirically to be capable of generating useful measurement data for each country and to have strong support across participating countries.

Test items for the student and school questionnaires are developed using similar processes to those used in developing the cognitive test items. Elements of the process

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Latvia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Argentina</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Liechtenstein</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Australia</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Lithuania</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Austria</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Luxembourg</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Macedonia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Belgium</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Mexico</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Brazil</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Netherlands</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>New Zealand</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Canada</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Norway</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chile</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Peru</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>China (Shanghai)</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Poland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hong Kong - China</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Portugal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Macao - China</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Qatar</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chinese Taipei</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Republic of Korea</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Colombia</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Romania</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Croatia</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Russian Federation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Scotland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Denmark</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Montenegro</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Serbia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Estonia</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Singapore</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Finland</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Slovak Republic</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>France</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Slovenia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Germany</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Greece</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Sweden</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hungary</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Switzerland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Iceland</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Thailand</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Indonesia</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Trinidad and Tobago</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ireland</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Tunisia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Israel</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Turkey</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Italy</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>United Kingdom</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Japan</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>United States of America</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jordan</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Uruguay</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>✓ ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

1. Includes participation in PISA plus (OECD 2003a), which was a replication of PISA 2000 run in 2001
2. The list of 2009 participants is provisional
include the work of professional questionnaire item writers, extensive consultation with relevant national experts in each participating country, use of cognitive laboratory procedures, and extensive field testing of items proposed for inclusion in the main survey.

**Sampling and Test Design**

PISA tests are administered to a random sample of students in each participating country. A two-stage sampling process is used. In the first stage, schools containing 15-year-olds are sampled with probability proportional to the number of eligible students. A minimum of 150 schools from each country is chosen. In the second stage, for each sampled school, about 30 students are randomly selected from those eligible. Hence a minimum of 4,500 students per country is selected to participate in PISA.

Each student is assigned a test booklet at random. In 2003 and 2006 PISA used a balanced incomplete block test design, involving 13 different booklets, each containing four blocks of test items. The blocks were populated with 13 “clusters” of items—7 clusters of items from the major domain and 3 from the remaining assessment domains. The clusters were rotated across booklets in such a way that each cluster appeared exactly once in each of the four block positions across the 13 booklets. This design enabled inclusion of a sufficient volume of material to guarantee coverage of the test domains and ensured that any item position effects were neutralized.

The approaches to school and student sampling and to assignment of material to the sampled students support the production of efficient estimates of PISA performance at the aggregated level of country and of major subgroups. PISA is not primarily designed for generating individual scores or ability estimates.

**Implementation in Participating Countries**

PISA tests are administered to students in their normal language of instruction. This means that the test instruments and questionnaires, which are generated centrally by the PISA consortium and dispatched as “source versions” to participating countries in both English and French, must be translated into the languages in common use in schools in each participating country. For PISA 2006, the 57 participating countries tested in more than 40 different languages in some 80 distinct national versions. Moreover, each national version of the PISA instruments should be equivalent to the source versions so that it can generate performance data that are comparable from country to country. This is a major technical challenge, and the translation and translation verification procedures form an important part of the work of participating countries and of the international consortium.

Operational procedures used for project implementation in each participating country, and in each test administration in each sampled school across all participating countries, are also as standard and common as possible. Project implementation is guided by a set of operations manuals that include a national project manager manual, a sampling manual, translation and verification guidelines, test administration procedures manuals, and a data management manual. The international consortium conducts training for NPMs, and sometimes other national center staff, in all aspects.
of project implementation at meetings that occur at critical points throughout each survey cycle.

The consortium also monitors project implementation using a variety of quality assurance procedures that include online mechanisms for collecting key data and monitoring implementation timelines, conducting interviews with national center staff, observing a sample of test administration sessions, collecting written reports, and regular e-mail and telephone communication. These quality monitoring activities have two main purposes. First, they give the consortium information about any implementation difficulties national centers may be experiencing so that support can be provided to overcome any problems before the integrity of PISA results is threatened. Second, they provide hard data about issues that may affect data integrity and which can be used at the formal stage of data adjudication to judge whether the data from each participating country are fit for use.

Scaling of Data
The scaling of PISA data is another very important technical aspect of the study. The approaches to scaling rest on the following assumptions: there are underlying traits (or sets of traits) of interest, which are defined in the relevant assessment frameworks, and each of which forms a continuum, or scale; test items can be developed that tap each of these underlying traits by requiring tested students to employ the trait in responding to the test items; and the amount of the trait possessed by students can be estimated by observing how successful they are in responding to the test items. Typically we think of the continuum as a line—with more of the trait in one direction on the line and less of the trait in the other direction—along which test items can be placed according to the amount of the underlying trait they demand and along which students can also be located according to the amount of the underlying trait they possess.

PISA uses a form of the Rasch model in order to scale the student data to derive the various comparative measures that are produced and reported by the OECD. Essentially this model uses student responses to a set of test items to simultaneously derive estimates of the difficulty of test items and of the ability of the sampled students, enabling both test items and students to be located along an interval scale that is defined by the underlying trait being measured. This is illustrated in Figure 1.2. Details of the model used can be found in the PISA 2000 Technical Report (Adams and Wu 2002).

The model and scaling methods also permit observed student responses to test items, or more specifically the ability measures obtained from the responses, to be linked to background variables, such as gender and socioeconomic background. Outcomes of such analyses are estimates of population means and other statistics that enable comparisons among groups of students between and within PISA-sampled populations. Those comparisons, linking performance in one or more of the cognitive test domains with various background variables, enable policy-oriented researchers to investigate factors that may influence student performance. In addition, the use
of common test and questionnaire items across successive survey administrations enables PISA data to be used to monitor change over time in outcome variables and in relationships among outcome and background variables.

**Reporting PISA Results**

In reporting the literacy outcomes within the major assessment domain, emphasis is placed on the profile of student results in each country in relation to the scales and subscales derived from the relevant framework. For example, in the case of scientific literacy, PISA 2006 results are reported for an overall science literacy scale, having a mean of 500 and a standard deviation of 100. The results are also reported separately for subscales that are based on the three competencies defined in the science framework. Central to the profile of student results is a set of descriptions of what students located at various points along the science literacy scale would typically be able to do,
As a first step in developing proficiency descriptions, the continuum of increasing scientific literacy was divided into five levels (bands), each of equal width, and two unbounded regions, one at each end of the continuum. The level definitions for PISA 2006 science are given in Figure 1.3. Information about the items in each level was then used to develop descriptions characterizing typical student performance at each level. Six proficiency descriptions were thus developed, with Level 6 being the highest level of proficiency. Student performance below Level 1 could not be reliably described because of the small number of PISA science items located in this region of the scale.

Figure 1.4 shows what it means for a student to be at one of Levels 1 to 5 in science. Each of these levels has a width of 74.6 science scale points. Two general principles common to all three PISA literacy scales are involved: (1) Students are located on the scale opposite the item(s) that they have a 62% chance of getting correct, and (2) “being at a level” means a student would be expected to get at least half of the items correct in a hypothetical test composed of questions spread uniformly across that level.

The figure shows that a student near the bottom of the level would have a 62% chance of getting an item at the bottom of the level correct, and a 42% chance of getting an item at the top of the level correct. A student at the top of the level has a 62% chance of getting an item at the top of the level correct, and a 78% chance of getting an item at the bottom of the level correct. Note that this latter student would not be expected to get 50% of the items correct on a test composed of items from the next higher level.
TABLE 1.2

WHAT IS PISA?

<table>
<thead>
<tr>
<th>Name</th>
<th>PISA is an acronym for the Programme for International Student Assessment. It is an international comparative survey program of the OECD.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Cycle</td>
<td>The assessment takes place every three years. PISA assesses reading literacy, mathematical literacy, and scientific literacy. Each survey cycle looks in depth at a major domain to which two-thirds of testing time is devoted; a summary profile of skills is provided for the other domains.</td>
</tr>
<tr>
<td>Basics</td>
<td>A standardized assessment that is jointly developed by participating countries and administered to 15-year-olds. Implemented in 43 countries in the first survey (32 in 2000 and 11 in 2002), 41 countries in the second survey (2003), and 57 in the third survey (2006). Implementation plans are under way for 65 countries in the fourth survey (2009). Typically administered to between 4,500 and 10,000 randomly selected students in each country. In 2006, approximately 500,000 students were assessed.</td>
</tr>
<tr>
<td>Content</td>
<td>PISA 2006 covers the domains of reading, mathematical, and scientific literacy with emphasis on important knowledge and skills needed in adult life. The major emphasis in PISA 2006 was on scientific competencies.</td>
</tr>
<tr>
<td>Methods</td>
<td>To date, paper-and-pencil tests have been used for the main assessment, with each student doing a two-hour test. Optional computer-developed test was available in 2006. Test items are a mixture of multiple-choice items and questions requiring students to construct their own responses. The items are organized in groups based on stimulus material setting out a real-life situation. About 390 minutes of test items was covered in 2006, with different students taking different combinations of test items. Students answer a background questionnaire, which takes 30 minutes to complete, providing information about themselves, their homes, and their school environment. School staff answer a 20-minute questionnaire about their school.</td>
</tr>
<tr>
<td>Outcomes</td>
<td>A basic profile of knowledge and skills among 15-year-old students. Contextual indicators relating results to student and school characteristics, with emphasis in 2006 placed on assessing students’ attitudes toward science. Trend indicators showing how results change over time. A valuable knowledge base for policy analysis and research.</td>
</tr>
</tbody>
</table>
The proficiency descriptions provide a clear picture of the way students are able to draw on the various scientific competencies described in the framework. They describe growth in scientific literacy in relation to an increasing student capacity to demonstrate and draw on those competencies. The main international OECD report of PISA 2006 outcomes (OECD 2007a) places significant emphasis on the relative proportions of students in each country performing at various levels along the scientific literacy continuum. Level 2 is described as the baseline level of scientific literacy, the level at which students begin to demonstrate the science competencies that will enable them to participate actively in life situations related to science and technology.

In Conclusion
Science teachers should be aware of the complexity, preparation, thoroughness, and standards that underpin an international assessment such as PISA 2006. This chapter has introduced some of the essential features of PISA, many of which are discussed in more detail in subsequent chapters of this book. Further information about PISA is available from the OECD website and at https://mypisa.acer.edu.au. A brief summary of PISA, adapted from OECD (2006), is presented in Table 1.2.

References

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACER. See Australia, Council for Educational Research</td>
<td></td>
</tr>
<tr>
<td>Achieving Scientific Literacy: From Purposes to Practices</td>
<td>228</td>
</tr>
<tr>
<td>Acid rain, 30, 72, 72–75, 251, 270–274</td>
<td></td>
</tr>
<tr>
<td>Adams, Ray, 5</td>
<td></td>
</tr>
<tr>
<td>Airbags safety, student interest in, 158</td>
<td></td>
</tr>
<tr>
<td>Alternative framework paradigm, science education, 103–104</td>
<td></td>
</tr>
<tr>
<td>Analytic approach, 124–125</td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
</tr>
<tr>
<td>interest in learning science, 152</td>
<td></td>
</tr>
<tr>
<td>mean science scores, 231</td>
<td></td>
</tr>
<tr>
<td>proficiency levels, 53</td>
<td></td>
</tr>
<tr>
<td>student interest in science topics, 155–156</td>
<td></td>
</tr>
<tr>
<td>Assessment, teaching, alignment of, 133–138</td>
<td></td>
</tr>
<tr>
<td>coherence among teaching goals, activities, 136–138</td>
<td></td>
</tr>
<tr>
<td>goals of, 133–134</td>
<td></td>
</tr>
<tr>
<td>science framework, 134–135</td>
<td></td>
</tr>
<tr>
<td>steps of evaluation process, 137</td>
<td></td>
</tr>
<tr>
<td>teacher, student ideas, gaps between, 136</td>
<td></td>
</tr>
<tr>
<td>test questions, 134–135</td>
<td></td>
</tr>
<tr>
<td>Assessment framework, 15–26, 19, 25–26</td>
<td></td>
</tr>
<tr>
<td>assessment, 18–19</td>
<td></td>
</tr>
<tr>
<td>attitude assessment, 24–25</td>
<td></td>
</tr>
<tr>
<td>competencies, 20–22, 21</td>
<td></td>
</tr>
<tr>
<td>contexts for PISA Science Assessment, 19–20, 20</td>
<td></td>
</tr>
<tr>
<td>defining, 16–18</td>
<td></td>
</tr>
<tr>
<td>explaining scientific phenomena, 21</td>
<td></td>
</tr>
<tr>
<td>identifying scientific issues, 21</td>
<td></td>
</tr>
<tr>
<td>knowledge of science categories, 23, 23–24, 24</td>
<td></td>
</tr>
<tr>
<td>scientific knowledge, 22–24</td>
<td></td>
</tr>
<tr>
<td>using scientific evidence, 22</td>
<td></td>
</tr>
<tr>
<td>Assessment of attitudes, 139–147</td>
<td></td>
</tr>
<tr>
<td>Assessment of competency, 215–224</td>
<td></td>
</tr>
<tr>
<td>performance-based assessment, 219–220</td>
<td></td>
</tr>
<tr>
<td>rubrics as tool for assessing performance-based tasks, 220–222</td>
<td></td>
</tr>
<tr>
<td>variety of assessments, need for, 222–223</td>
<td></td>
</tr>
<tr>
<td>Astronomy, percentage of students interested in, 155</td>
<td></td>
</tr>
<tr>
<td>Atomic/molecular model, 209</td>
<td></td>
</tr>
<tr>
<td>Attitude assessment, 139–147</td>
<td></td>
</tr>
<tr>
<td>aspects of attitudes toward science, 140</td>
<td></td>
</tr>
<tr>
<td>indices of attitudes toward science, 141</td>
<td></td>
</tr>
<tr>
<td>interest in science, 141–143</td>
<td></td>
</tr>
<tr>
<td>mean score of interest in science, 143</td>
<td></td>
</tr>
<tr>
<td>purpose of learning science, 144–145</td>
<td></td>
</tr>
<tr>
<td>reflective citizen, acting as, 17, 145–146</td>
<td></td>
</tr>
<tr>
<td>scales of attitudes toward science, 141</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td></td>
</tr>
<tr>
<td>achievement scores, 167</td>
<td></td>
</tr>
<tr>
<td>Council for Educational Research, 5</td>
<td></td>
</tr>
<tr>
<td>experimentation, participation in, 171</td>
<td></td>
</tr>
<tr>
<td>interactive activities, 170</td>
<td></td>
</tr>
<tr>
<td>interest in learning science, 152</td>
<td></td>
</tr>
<tr>
<td>investigation activity, participation in, 173</td>
<td></td>
</tr>
<tr>
<td>knowledge profile, 63</td>
<td></td>
</tr>
<tr>
<td>mean science scores, 231</td>
<td></td>
</tr>
<tr>
<td>proficiency levels, 53</td>
<td></td>
</tr>
<tr>
<td>science performance, 166</td>
<td></td>
</tr>
<tr>
<td>scientific literacy, 113, 127</td>
<td></td>
</tr>
<tr>
<td>student interest in science topics, 155–156</td>
<td></td>
</tr>
<tr>
<td>teacher explanation of science, real world linkage, 174</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td></td>
</tr>
<tr>
<td>knowledge profile, 63</td>
<td></td>
</tr>
<tr>
<td>mean science scores, 231</td>
<td></td>
</tr>
<tr>
<td>proficiency levels, 53</td>
<td></td>
</tr>
<tr>
<td>scientific literacy, 127</td>
<td></td>
</tr>
<tr>
<td>Azerbaijan</td>
<td></td>
</tr>
<tr>
<td>mean science scores, 231</td>
<td></td>
</tr>
<tr>
<td>proficiency levels, 53</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td></td>
</tr>
<tr>
<td>knowledge profile, 63</td>
<td></td>
</tr>
<tr>
<td>mean science scores, 231</td>
<td></td>
</tr>
</tbody>
</table>
proficiency levels, 53
scientific literacy, 127

Biological evolution, 105–107, 209

Brazil
interest in learning science, 152
mean science scores, 231
proficiency levels, 53
student interest in science topics, 155–156

Breathing rate during physical exercise, student interest in, 158

Broad science topics, percentage of students interested in, 155

Bulgaria
mean science scores, 231
proficiency levels, 53

Bybee, Rodger, 228

Canada, 7, 91–99
interest in learning science, 152
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 113, 127
student interest in science topics, 155–156
teaching activity, 120

Cancer of skin, sunscreen use, 73, 73, 251, 265–269
student interest in, 158
Carbon chemical cycle, 209

Cells, as basic units of living things, 209
Chemical composition of agricultural fertilizers, student interest, 159

Chemical cycles, 209
Chemistry, percentage of students interested in, 155

China
mean science scores, 231
proficiency levels, 53

Citizen, reflective, 17
acting as, 17, 145–146

CITOGroup, 5

Coding guides, 29–30, 30

Coherence among teaching goals, activities, 136–138

Colombia
interest in learning science, 152
mean science scores, 231
proficiency levels, 53
student interest in science topics, 155–156

Competency assessment, 215–224
performance-based assessment, 219–220
rubrics as tool for assessing performance-based tasks, 220–222
variety of assessments, need for, 222–223

Conant, James Bryant, 227

Contexts for scientific, technological literacy, 243

Corn, genetically modified, 74, 74, 251, 279–280

Coverage of science content in Third International Mathematics and Science Study, PISA, 65

Croatia
mean science scores, 231
proficiency levels, 53

Crops, genetically modified, 74, 251, 279–280

Cultural appropriateness, 41–42

Curriculum, 241–244
design, 203–214, 211
models, 205, 207
pilot trial outcomes, 212–213

Czech Republic, 7, 70–75
achievement scores, 167
experimentation, participation in, 171
interactive activities, 170
investigation activity, participation in, 173
knowledge profile, 63
mean science scores, 231
proficiency levels, 53

Disease transmission, student interest in, 158

Determinants of student interest, 156–157

Digestion, pioneers in study of, student interest, 159

Disease transmission, student interest in, 158

Earth, 23, 85
structure, evolution of, 209

Education sources, transfers, 209

English-speaking countries, 106

Estonia
mean science scores, 231
proficiency levels, 53

Evaluation framework, 15–26, 19, 25–26
assessment, 18–19
attitude assessment, 24–25
competencies, 20–22, 21
contexts for PISA Science Assessment, 19–20, 20

Coherence among teaching goals, activities, 136–138

Colombia
interest in learning science, 152
mean science scores, 231
proficiency levels, 53
student interest in science topics, 155–156

Competency assessment, 215–224
performance-based assessment, 219–220
rubrics as tool for assessing performance-based tasks, 220–222
variety of assessments, need for, 222–223

Conant, James Bryant, 227

Contexts for scientific, technological literacy, 243

Corn, genetically modified, 74, 74, 251, 279–280

Coverage of science content in Third International Mathematics and Science Study, PISA, 65

Croatia
mean science scores, 231
proficiency levels, 53

Crops, genetically modified, 74, 74, 251, 279–280

Cultural appropriateness, 41–42

Curriculum, 241–244
design, 203–214, 211
models, 205, 207
pilot trial outcomes, 212–213

Czech Republic, 7, 70–75
achievement scores, 167
experimentation, participation in, 171
interactive activities, 170
investigation activity, participation in, 173
knowledge profile, 63
mean science scores, 231
proficiency levels, 53

Disease transmission, student interest in, 158

Determinants of student interest, 156–157

Digestion, pioneers in study of, student interest, 159

Disease transmission, student interest in, 158

Earth, 23, 85
structure, evolution of, 209

Education sources, transfers, 209

English-speaking countries, 106

Estonia
mean science scores, 231
proficiency levels, 53

Evaluation framework, 15–26, 19, 25–26
assessment, 18–19
attitude assessment, 24–25
competencies, 20–22, 21
contexts for PISA Science Assessment, 19–20, 20
defining, 16–18
explaining scientific phenomena, 21
identifying scientific issues, 21
knowledge of science categories, 23, 24
scientific knowledge, 22–24
using scientific evidence, 22
Evaluation of attitudes, 139–147
Evaluation of competency, 215–224
performance-based assessment, 219–220
rubrics as tool for assessing performance-based tasks, 220–222
variety of assessments, need for, 222–223
Evaluation of teaching, alignment of, 133–138
cohesion among teaching goals, activities, 136–138
goals of, 133–134
science framework, 134–135
steps of evaluation process, 137
teacher, student ideas, gaps between, 136
test questions, 134–135
Evans, Robert, 229
Evidence, scientific, use of, 21–22, 199–200
Evolution, 105–107, 209–211
Exercise, 251, 275–278
breathing rate during, student interest in, 158
muscular effects, student interest in, 158
Experimental design, percentage of students interested in, 155
Explanation of scientific phenomena, 21, 197–199

Fertilizers effects, experimental design, student interest, 159
Field trials, 34–35
analyses, 34
cluster rotation design, 35
international coder support, 34
national field trial reviews, 34
Finnish students, 7, 79–90, 83, 106
assessment, 80–82
content areas, 85
curriculum, 80–82
equal comprehensive schools, 80
goals, 85
interest in learning science, 152
knowledge profile, 63
learning environments, 85–86
learning materials, 85–86
mean science scores, 231
physics, chemistry, biology teachers, 86
proficiency levels, 53
science curriculum, 82–85
scientific competencies, 84–85
scientific literacy, 113, 127
society, science education and, 87–88
student interest in science topics, 155–156
Formulation of policies for scientific literacy, 239
Framework for assessment, 15–26, 19, 25–26
assessment, 18–19
attitude assessment, 24–25
competencies, 20–22
contexts for PISA Science Assessment, 19–20
defining, 16–18
explaining scientific phenomena, 21
knowledge of science categories, 23, 24
scientific knowledge, 22–24
using scientific evidence, 22
Framework for science curricula, 242
France, 7, 70–75, 106–107
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
reforms in teaching of science, 75
scientific literacy, 127
student competency levels, 71
student performance, 75–76
Fuels used in portable stoves, student interest, 159

Germany
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 127
GM crops. See Genetically modified crops
Grand Canyon, 251, 261–264
Greece
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 127
Greenhouse effect, 74–75, 75–76, 251, 254–260

Handbook of Research on Science Education, 229
Hearing sensitivity, student interest in, 158
High achievement, 123–132, 182–185
combining learning experiences for, 128–130
learning experiences with, 126–128
Hong Kong
mean science scores, 231
proficiency levels, 53
Human biology, percentage of students interested in, 155
Hungary
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 127
Iceland, 7, 106
mean science scores, 231
proficiency levels, 53
scientific literacy, 127
Identification of scientific issues, 21, 196–197
Immigrant students, 92–93
Immune system, student interest in, 158
Indices of attitudes toward science, 141
Indonesia
mean science scores, 231
proficiency levels, 53
Inheritance, gene theory of, 209
International comparisons, 47–176
Ireland
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 127
Israel
  mean science scores, 231
  proficiency levels, 53
Italy
  knowledge profile, 63
  mean science scores, 231
  proficiency levels, 53
  scientific literacy, 127
Item-specific information, utilization of, 102–103
Japan
  achievement scores, 167
  experimentation, participation in, 171
  interactive activities, 170
  interest in learning science, 152
  investigation activity, participation in, 173
  knowledge profile, 63
  mean science scores, 231
  proficiency levels, 53
  science performance, 166
  scientific literacy, 113, 127
  student interest in science topics, 155–156
  teacher explanation of science, real world linkage, 174
Jordan
  mean science scores, 231
  proficiency levels, 53
Kemp, Andrew, 229
Koballa, Thomas, 229
Korea, Republic of
  knowledge profile, 63
  proficiency levels, 53
  scientific literacy, 113, 127
Kyrgyz Republic
  mean science scores, 231
  proficiency levels, 53
Laos
  mean science scores, 231
  proficiency levels, 53
Lithuania
  mean science scores, 231
  proficiency levels, 53
Loud music, damage to hearing, student interest in, 158
Luxembourg
  knowledge profile, 63
  mean science scores, 231
  proficiency levels, 53
  scientific literacy, 127
Macao
  mean science scores, 231
  proficiency levels, 53
Main survey test design, 35–37
Malleus, Pierre, 75
Mean score of interest in science, 143
Mexico
  interest in learning science, 152
  knowledge profile, 63
  mean science scores, 231
  proficiency levels, 53
  scientific literacy, 127
  student interest in science topics, 155–156
Miller, Robin, 229
Molecular structures of plastics, student interest, 159
Montenegro
  mean science scores, 231
  proficiency levels, 53
Motivation, 43–44
Multiple-choice items, 29
Music, damage to hearing, student interest in, 158
NAEP. See National Assessment of Educational Progress
National Assessment of Educational Progress, 238
National Institute for Educational Measurement, 5
National item reviews, 32–33
Natural selection, 209
Netherlands, 7, 107
  achievement scores, 167
  experimentation, participation in, 171
  interactive activities, 170
  interest in learning science, 152
  investigation activity, participation in, 173
  knowledge profile, 63
mean science scores, 231
  proficiency levels, 53
  science performance, 166
  scientific literacy, 127
  student interest in science topics, 155–156
  teacher explanation of science, real world linkage, 174
New Zealand
  interest in learning science, 152
  knowledge profile, 63
  mean science scores, 231
  proficiency levels, 53
  scientific literacy, 127
OECD countries. See Organisation for Economic Cooperation and Development countries
Organisation for Economic Cooperation and Development countries, 3, 121
  knowledge profiles, 63
  scientific literacy scores, 127
  secretariat oversight, 4
  similarities in knowledge profiles, 63
  website, 3
Osborne, Jonathan, 229–230
Perceptions of science teaching, 119–121
Performance-based assessment, 219–220
Physical exercise, 251, 275–278
  breathing rate during, student interest in, 158
  muscular effects, student interest in, 158
Physics, percentage of students interested in, 155
PISA, 3–13
  assessment framework, 15–26
  attitudes, 139–147, 149–162
INDEX

mean science scores, 231
proficiency levels, 53
scientific literacy, 127

Switzerland
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 127

Taipei
mean science scores, 231
proficiency levels, 53

Test development, 6–8
adaptation, 33–34
cluster rotation design, 35
coding guides, 29–30, 30
design, 27–38
dual-source versions, 33
field trial, 34–35
field trial analyses, 34
international coder support, 34
international pilot testing, 32
item development, 30–33
item types, 28–29
main survey items, 36–37
main survey test design, 35–36
multiple-choice items, 29
national field trial reviews, 34
national item reviews, 32–33
national submissions, 31
process of, 32–33
test items, 27–30
translation, 33–34
unit structure, 27–28

Testing, student attitudes, 149–162, 151
broad science topics, percentage of students interested in, 155
determinants of student interest, 156–157
embedded attitudinal test items, 150–151
gender differences, 153–154
gender preferences, 157–160
general interest in learning science, 154
interest in learning science topics scale, 151–154
interest in specific topics, 154–160
specific science topics, percentage of students interested in, 156

Thailand
mean science scores, 231
proficiency levels, 53
Theory of evolution, 105–107, 209
Third International Mathematics and Science Study, 163–176, 166
PISA, compared, 65–66
TIMSS. See Third International Mathematics and Science Study

Tunisia
mean science scores, 231
proficiency levels, 53

Turkey
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 127

Twenty First Century Science, 229

United Kingdom
knowledge profile, 63
mean science scores, 231
proficiency levels, 53
scientific literacy, 113, 127

United States, 7, 70–75, 111–115, 179–186
achievement scores, 167
experimentation, participation in, 171
high-achieving countries, lessons from, 182–185
interactive activities, 170
interest in learning science, 152
investigation activity, participation in, 173
knowledge profile, 63
loss of educational advantage, 180
mean science scores, 231
PISA performance, 180
proficiency levels, 53
science performance, 166
scientific literacy, 113, 127
student competency levels, 71
student interest in science topics, 155–156
teacher explanation of science, real world linkage, 174
teaching activity, 120

Universe, structure, evolution of, 209
Uruguay
interest in learning science, 152
mean science scores, 231
proficiency levels, 53
student interest in science topics, 155–156

Variety of assessments, need for, 222–223

Water, atomic structure, student interest, 159

Website of Organisation for Economic Cooperation and Development, 3
Weights in assessment framework, 3
Wood, atomic structure, student interest, 159