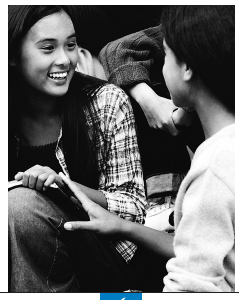


1

Programme for International Student Assessment: An Overview

Participation 21
Features of PISA 22
Managing and implementing PISA 23
Organisation of this report 23

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

The OECD Programme for International Student Assessment (PISA) is a collaborative effort among OECD member countries to measure how well 15-year-old students approaching the end of compulsory schooling are prepared to meet the challenges of today's knowledge societies. The assessment is forward-looking: rather than focusing on the extent to which these students have mastered a specific school curriculum, it looks at their ability to use their knowledge and skills to meet real-life challenges. This orientation reflects a change in curricular goals and objectives, which are increasingly concerned with what students can do with what they learn at school.

PISA surveys take place every three years. The first survey took place in 2000 (followed by a further 11 countries in 2001), the second in 2003 and the third in 2006; the results of these surveys have been published in a series of reports (OECD, 2001, 2003, 2004, 2007) and a wide range of thematic and technical reports. The next survey will occur in 2009. For each assessment, one of the three areas (science, reading and mathematics) is chosen as the major domain and given greater emphasis. The remaining two areas, the minor domains, are assessed less thoroughly. In 2000 the major domain was reading; in 2003 it was mathematics and in 2006 it was science.

PISA is an age-based survey, assessing 15-year-old students in school in grade seven or higher. These students are approaching the end of compulsory schooling in most participating countries, and school enrolment at this level is close to universal in almost all OECD countries.

The PISA assessments take a literacy perspective, which focuses on the extent to which students can apply the knowledge and skills they have learned and practised at school when confronted with situations and challenges for which that knowledge may be relevant. That is, PISA assesses the extent to which students can use their reading skills to understand and interpret the various kinds of written material that they are likely to meet as they negotiate their daily lives; the extent to which students can use their mathematical knowledge and skills to solve various kinds of numerical and spatial 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. The PISA 2006 domain definitions are fully articulated in *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (OECD, 2006).

PISA also allows for the assessment of additional cross-curricular competencies from time to time as participating countries see fit. For example, in PISA 2003, an assessment of general problem-solving competencies was included. Further, PISA uses student questionnaires to collect information from students on various aspects of their home, family and school background, and school questionnaires to collect information from schools about various aspects of organisation and educational provision in schools. In PISA 2006 a number of countries[1] also administered a parent questionnaire to the parents of the students participating in PISA.

Using the data from these questionnaires, analyses linking contextual information with student achievement could address:

- Differences between countries in the relationships between student-level factors (such as gender and social background) and achievement;
- Differences in the relationships between school-level factors and achievement across countries;
- Differences in the proportion of variation in achievement between (rather than within) schools, and differences in this value across countries;
- Differences between countries in the extent to which schools moderate or increase the effects of individual-level student factors and student achievement;

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

- Differences in education systems and national context that are related to differences in student achievement across countries;
- Through links to PISA 2000 and PISA 2003, changes in any or all of these relationships over time.

Through the collection of such information at the student and school level on a cross-nationally comparable basis, PISA adds significantly to the knowledge base that was previously available from national official statistics, such as aggregate national statistics on the educational programmes completed and the qualifications obtained by individuals.

PARTICIPATION

The first PISA survey was conducted in 2000 in 32 countries (including 28 OECD member countries) using written tasks answered in schools under independently supervised test conditions. Another 11 countries completed the same assessment in 2001. PISA 2000 surveyed reading, mathematical and scientific literacy, with a primary focus on reading. The second PISA survey, conducted in 2003 in 41 countries, assessed reading, mathematical and scientific literacy, and problem solving with a primary focus on mathematical literacy. The third survey covered reading, mathematical and scientific literacy, with a primary focus on scientific literacy, and was conducted in 2006 in 57 countries. In some countries it was decided to carry out detailed analysis of some regions. In these 24 sub-national regions sufficient data were collected and quality control mechanisms implemented at a level that would permit OECD endorsement of their results. The participants in PISA 2006 are listed in Table 1.1. This report is concerned with the technical aspects of PISA 2006.

Table 1.1
PISA 2006 participants

OECD countries	Partner countries/economies
Australia	Argentina
Austria	Azerbaijan
Belgium	Brazil
Canada	Bulgaria
Czech Republic	Chile
Denmark	Colombia
Finland	Croatia
France	Estonia
Germany	Hong Kong-China
Greece	Indonesia
Hungary	Israel
Iceland	Jordan
Ireland	Kyrgyzstan
Italy	Latvia
Japan	Liechtenstein
Korea	Lithuania
Luxembourg	Macao-China
Mexico	Montenegro
Netherlands	Qatar
New Zealand	Romania
Norway	Russian Federation
Poland	Serbia
Portugal	Slovenia
Slovak Republic	Chinese Taipei
Spain	Thailand
Sweden	Tunisia
Switzerland	Uruguay
Turkey	
United Kingdom	
United States	

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

FEATURES OF PISA

The technical characteristics of the PISA survey involve a number of different challenges:

- The design of the test and the features incorporated into the test developed for PISA are critical;
- The sampling design, including both the school sampling and the student sampling requirements and procedures;
- The multilingual nature of the test, which involves rules and procedures designed to guarantee the equivalence of the different language versions used within and between participating countries, and taking into account the diverse cultural contexts of those countries;
- Various operational procedures, including test administration arrangements, data capture and processing and quality assurance mechanisms designed to ensure the generation of comparable data from all countries;
- Scaling and analysis of the data and their subsequent reporting. PISA employs scaling models based on item response theory (IRT) methodologies. The described proficiency scales, which are the basic tool in reporting PISA outcomes, are derived using IRT analysis.

Box 1.1 Core features of PISA 2006

Sample size

- Nearly 400,000 students, representing almost 20 million 15-year-olds enrolled in the schools of the 57 participating countries and economies, were assessed in 2006.

Content

- PISA 2006 covered three domains: reading, mathematics and science.
- PISA 2006 looked at young people's ability to use their knowledge and skills in order to meet real-life challenges rather than how well they had mastered a specific school curriculum. The emphasis was placed on the mastery of processes, the understanding of concepts, and the ability to function in various situations within each domain.

Methods

- PISA 2006 used pencil-and-paper assessments, lasting two hours for each student.
- PISA 2006 used both multiple-choice items and questions requiring students to construct their own answers. Items were typically organised in units based on a passage describing a real-life situation.
- A total of six and a half hours of assessment items was included, with different students taking different combinations of the assessment items.
- Students answered a background questionnaire that took about 30 minutes to complete and, as part of an international option, completed questionnaires on their educational careers as well as familiarity with computers.
- School principals completed questionnaires about their schools.

Outcomes

- A profile of knowledge and skills among 15-year-olds.
- Contextual indicators relating results to student and school characteristics.
- A knowledge base for policy analysis and research.
- Trend indicators showing how results change over time.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

This report describes the above-mentioned methodologies as they have been implemented in PISA 2006. Further, it describes the quality assurance procedures that have enabled PISA to provide high quality data to support policy formation and review. Box 1.1 provides an overview of the central design elements of PISA 2006.

The ambitious goals of PISA come at a cost: PISA is both resource intensive and methodologically complex, requiring intensive collaboration among many stakeholders. The successful implementation of PISA depends on the use, and sometimes further development, of state-of-the-art methodologies.

Quality within each of these areas is defined, monitored and assured through the use of a set of technical standards. These standards have been endorsed by the PISA Governing Board, and they form the backbone of implementation in each participating country and of quality assurance across the project.

MANAGING AND IMPLEMENTING PISA

The design and implementation of PISA for the 2000, 2003 and 2006 data collections has been the responsibility of an international consortium led by the Australian Council for Educational Research (ACER) with Ray Adams as international project director. The other partners in this consortium have been the National Institute for Educational Measurement (Cito Group) in the Netherlands, Unité d'analyse des systèmes et des pratiques d'enseignement (aSPe) at Université de Liège in Belgium, Westat and the Educational Testing Service (ETS) in the United States and the National Institute for Educational Research (NIER) in Japan. Appendix 8 lists the consortium staff and consultants who have made significant contributions to the development and implementation of the project.

The consortium implements PISA within a framework established by the PISA Governing Board (PGB) which includes representation from all participating countries at senior policy levels. The PGB established policy priorities and standards for developing indicators, for establishing assessment instruments, and for reporting results. Experts from participating countries served on working groups linking the programme policy objectives with the best internationally available technical expertise in the three assessment areas. These expert groups were referred to as Subject Matter Expert Groups (SMEGs) (*see* Appendix 8 for members). By participating in these expert groups and regularly reviewing outcomes of the groups' meetings, countries ensured that the instruments were internationally valid and that they took into account the cultural and educational contexts of the different OECD member countries, that the assessment materials had strong measurement potential, and that the instruments emphasised authenticity and educational validity.

Each of the participating countries appointed a National Project Manager (NPM), to implement PISA nationally. The NPM ensured that internationally agreed common technical and administrative procedures were employed. These managers played a vital role in developing and validating the international assessment instruments and ensured that PISA implementation was of high quality. The NPMs also contributed to the verification and evaluation of the survey results, analyses and reports.

The OECD Secretariat had overall responsibility for managing the programme. It monitored its implementation on a day-to-day basis, served as the secretariat for the PGB, fostered consensus building between the countries involved, and served as the interlocutor between the PGB and the international consortium.

ORGANISATION OF THIS REPORT

This technical report is designed to describe the technical aspects of the project at a sufficient level of detail to enable review and, potentially, replication of the implemented procedures and technical solutions to

problems. It, therefore, does not report the results of PISA 2006 which have been published in *PISA 2006: Science Competencies for Tomorrow's World* (OECD, 2007). A bibliography of other PISA related reports is included in Appendix 9.

There are five sections in this report:

- *Section One – Instrument design* (Chapters 1-4): Describes the design and development of both the questionnaires and achievement tests;
- *Section Two – Operations* (Chapters 5-7): Gives details of the operational procedures for the sampling and population definitions, test administration procedures, quality monitoring and assurance procedures for test administration and national centre operations, and instrument translation;
- *Section Three – Data processing* (Chapters 8-10): Covers the methods used in data cleaning and preparation, including the methods for weighting and variance estimation, scaling methods, methods for examining inter-rater variation and the data cleaning steps;
- *Section Four – Quality indicators and outcomes* (Chapters 11-14): Covers the results of the scaling and weighting, reports response rates and related sampling outcomes and gives the outcomes of the inter-rater reliability studies. The last chapter in this section summarises the outcomes of the PISA 2006 data adjudication; that is, the overall analysis of data quality for each country;
- *Section Five – Scale construction and data products* (Chapters 15-18): Describes the construction of the PISA 2006 levels of proficiency and the construction and validation of questionnaire-related indices. The final chapter briefly describes the contents of the PISA 2006 database;
- *Appendices*: Detailed appendices of results pertaining to the chapters of the report are provided.

Notes

1. The PISA 2006 parent questionnaire was administered in Denmark, Germany, Iceland, Italy, Luxembourg, New Zealand, Poland, Portugal, Korea and Turkey, as well as in the partner countries/economies Bulgaria, Colombia, Croatia, Hong Kong-China, Macao-China and Qatar.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Reader's Guide

Country codes – the following country codes are used in this report:

OECD countries

AUS	Australia
AUT	Austria
BEL	Belgium
BEF	Belgium (French Community)
BEN	Belgium (Flemish Community)
CAN	Canada
CAE	Canada (English Community)
CAF	Canada (French Community)
CZE	Czech Republic
DNK	Denmark
FIN	Finland
FRA	France
DEU	Germany
GRC	Greece
HUN	Hungary
ISL	Iceland
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
LUX	Luxembourg
LXF	Luxembourg (French Community)
LXG	Luxembourg (German Community)
MEX	Mexico
NLD	Netherlands
NZL	New Zealand
NOR	Norway
POL	Poland
PRT	Portugal
SVK	Slovak Republic
ESP	Spain
ESB	Spain (Basque Community)
ESC	Spain (Catalonian Community)
ESS	Spain (Castillian Community)
SWE	Sweden
CHE	Switzerland
CHF	Switzerland (French Community)
CHG	Switzerland (German Community)
CHI	Switzerland (Italian Community)
TUR	Turkey
GBR	United Kingdom
IRL	Ireland
SCO	Scotland
USA	United States

Partner countries and economies

ARG	Argentina
AZE	Azerbaijan
BGR	Bulgaria
BRA	Brazil
CHL	Chile
COL	Colombia
EST	Estonia
HKG	Hong Kong-China
HRV	Croatia
IDN	Indonesia
JOR	Jordan
KGZ	Kyrgyztan
LIE	Liechtenstein
LTU	Lithuania
LVA	Latvia
LVL	Latvia (Latvian Community)
LVR	Latvia (Russian Community)
MAC	Macao-China
MNE	Montenegro
QAT	Qatar
ROU	Romania
RUS	Russian Federation
SRB	Serbia
SVN	Slovenia
TAP	Chinese Taipei
THA	Thailand
TUN	Tunisia
URY	Uruguay

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

References

Adams, R.J., Wilson, M. & **Wang, W.C.** (1997), The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement,* No. 21, pp. 1-23.

Adams, R.J., Wilson, M. R. & **Wu, M.L.** (1997), Multilevel item response models: An approach to errors in variables regression, *Journal of Educational and Behavioural Statistics,* No. 22 (1), pp. 46-75.

Adams, R.J. & **Wu, M.L.** (2002), *PISA 2000 Technical Report,* OECD, Paris.

Bollen, K.A. & **Long, S.J.** (1993) (eds.), *Testing Structural Equation Models*, Newbury Park: London.

Beaton, A.E. (1987), Implementing the new design: The NAEP 1983-84 technical report (Rep. No. 15-TR-20), Princeton, NJ: Educational Testing Service.

Buchmann, C. (2000), Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school? *Soc. Forces,* No. 78, pp. 1349-79.

Buchmann, C. (2002), Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges, in Porter, A.C. and Gamoran, A. (eds.). *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-97), Washington, DC: National Academy Press.

Creemers, B.P.M. (1994), *The Effective Classroom*, London: Cassell.

Cochran, W.G. (1977), *Sampling techniques,* third edition, New York, NY: John Wiley and Sons.

Ganzeboom, H.B.G., de Graaf, P.M. & **Treiman, D.J.** (1992), A standard international socio-economic index of occupational status, *Social Science Research*, No. 21, pp. 1-56.

Ganzeboom H.B. & **Treiman, D.J.** (1996), Internationally comparable measures of occupational status for the 1988 international standard classification of occupations, *Social Science Research*, No. 25, pp. 201-239.

Grisay, A. (2003), Translation procedures in OECD/PISA 2000 international assessment, *Language Testing,* No. 20 (2), pp. 225-240.

Hambleton, R.K., Swaminathan, H. & **Rogers, H.J.** (1991), *Fundamentals of item response theory*, Newbury Park, London, New Delhi: SAGE Publications.

Hambleton, R.K., Merenda, P.F. & **Spielberger, C.D.** (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment,* IEA Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.

Harkness, J.A., Van de Vijver, F.J.R. & **Mohler, P.Ph** (2003), *Cross-Cultural Survey Methods,* Wiley-Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey.

Harvey-Beavis, A. (2002), Student and School Questionnaire Development, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report,* (pp. 33-38), OECD, Paris.

International Labour Organisation (ILO) (1990), *International Standard Classification of Occupations: ISCO-88.* Geneva: International Labour Office.

Jöreskog, K.G. & **Sörbom, Dag** (1993), *LISREL 8 User's Reference Guide*, Chicago: SSI.

Judkins, D.R. (1990), Fay's Method of Variance Estimation, *Journal of Official Statistics*, No. 6 (3), pp. 223-239.

Kaplan, D. (2000), *Structural equation modeling: Foundation and extensions,* Thousand Oaks: SAGE Publications.

Keyfitz, N. (1951), Sampling with probabilities proportionate to science: Adjustment for changes in probabilities, *Journal of the American Statistical Association,* No. 46, American Statistical Association, Alexandria, pp. 105-109.

Kish, L. (1992), Weighting for Unequal, *Pi. Journal of Official Statistics*, No. 8 (2), pp. 183-200.

LISREL (1993), K.G. Jöreskog & D. Sörbom, [computer software], Lincolnwood, IL: Scientific Software International, Inc.

Lohr, S.L. (1999), *Sampling: Design and Analysis*, Duxberry: Pacific Grove.

Macaskill, G., Adams, R.J. & **Wu, M.L.** (1998), Scaling methodology and procedures for the mathematics and science literacy, advanced mathematics and physics scale, in M. Martin and D.L. Kelly, Editors, *Third International Mathematics and Science Study, technical report Volume 3: Implementation and analysis*, Boston College, Chestnut Hill, MA.

Masters, G.N. & **Wright, B.D.** (1997), The Partial Credit Model, in W.J. van der Linden, & R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (pp. 101-122), New York/Berlin/Heidelberg: Springer.

Mislevy, R.J. (1991), Randomization-based inference about latent variables from complex samples, *Psychometrika,* No. 56, pp. 177-196.

Mislevy, R.J., Beaton, A., Kaplan, B.A. & **Sheehan, K.** (1992), Estimating population characteristics from sparse matrix samples of item responses, *Journal of Educational Measurement,* No. 29 (2), pp. 133-161.

Mislevy, R.J. & **Sheehan, K.M.** (1987), Marginal estimation procedures, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report,* National Assessment of Educational Progress, Educational Testing Service, Princeton, pp. 293-360.

Mislevy, R.J. & **Sheehan, K.M.** (1989), Information matrices in latent-variable models, *Journal of Educational Statistics,* No. 14, pp. 335-350.

Mislevy, R.J. & **Sheehan, K.M.** (1989), The role of collateral information about examinees in item parameter estimation, *Psychometrika,* No. 54, pp. 661-679.

Monseur, C. & **Berezner, A.** (2007), The Computation of Equating Errors in International Surveys in Education, *Journal of Applied Measurement,* No. 8 (3), 2007, pp. 323-335.

Monseur, C. (2005), An exploratory alternative approach for student non response weight adjustment, *Studies in Educational Evaluation,* No. 31 (2-3), pp. 129-144.

Muthen, B. & **L. Muthen** (1998), [computer software], *Mplus* Los Angeles, CA: Muthen & Muthen.

Muthen, B., du Toit, S.H.C. & **Spisic, D.** (1997), *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes,* unpublished manuscript.

OECD (1999), *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries,* OECD, Paris.

OECD (2003), *Literacy Skills for the World of Tomorrow: Further results from PISA 2000,* OECD, Paris.

OECD (2004), *Learning for Tomorrow's World – First Results from PISA 2003,* OECD, Paris.

OECD (2005), *Technical Report for the OECD Programme for International Student Assessment 2003,* OECD, Paris.

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006,* OECD, Paris.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World,* OECD, Paris.

PISA Consortium (2006), *PISA 2006 Main Study Data Management Manual, https://mypisa.acer.edu.au/images/mypisadoc/opmanual/pisa2006_data_management_manual.pdf*

Rasch, G. (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen: Nielsen & Lydiche.

Routitski A. & **Berezner, A.** (2006), Issues influencing the validity of cross-national comparisons of student performance. Data Entry Quality and Parameter Estimation. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Francisco, 7-11 April, *https://mypisa.acer.edu.au/images/mypisadoc/aera06routitsky_berezner.pdf*

Rust, K. (1985), Variance Estimation for Complex Estimators in Sample Surveys, *Journal of Official Statistics,* No. 1, pp. 381-397.

Rust, K.F. & **Rao, J.N.K.** (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Survey Methods in Medical Research,* No. 5, pp. 283-310.

Shao, J. (1996), Resampling Methods in Sample Surveys (with Discussion), *Statistics,* No. 27, pp. 203-254.

Särndal, C.-E., Swensson, B. & **Wretman, J.** (1992), *Model Assisted Survey Sampling,* New York: Springer-Verlag.

SAS® CALIS (1992), W. Hartmann [computer software], Cary, NC: SAS Institute Inc.

Scheerens, J. (1990), School effectiveness and the development of process indicators of school functioning, *School effectiveness and school improvement,* No. 1, pp. 61-80.

Scheerens, J. & **Bosker, R.J.** (1997), *The Foundations of School Effectiveness,* Oxford: Pergamon.

Schulz, W. (2002), Constructing and Validating the Questionnaire composites, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report,* OECD, Paris.

Schulz, W. (2004), Mapping Student Scores to Item Responses, in W. Schulz and H. Sibberns (eds.), *IEA Civic Education Study, Technical Report* (pp. 127-132), Amsterdam: IEA.

Schulz, W. (2006a), *Testing Parameter Invariance for Questionnaire Indices using Confirmatory Factor Analysis and Item Response Theory,* Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Schulz, W. (2006b), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003,* Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Thorndike, R.L. (1973), *Reading comprehension in fifteen countries,* New York, Wiley: and Stockholm: Almqvist & Wiksell.

Travers, K.J. & **Westbury, I.** (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula,* Oxford: Pergamon Press.

Travers, K.J., Garden R.A. & **Rosier, M.** (1989), Introduction to the Study, in Robitaille, D. A. and Garden, R. A. (eds), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics Curricula,* Oxford: Pergamon Press.

Verhelst, N. (2002), Coder and Marker Reliabilaity Studies, in R.J. Adams & M.L. Wu (eds.), *PISA 2000 Technical Report*. OECD, Paris.

Walberg, H.J. (1984), Improving the productivity of American schools, *Educational Leadership,* No. 41, pp. 19-27.

Walberg, H. (1986), Synthesis of research on teaching, in M. Wittrock (ed.), *Handbook of research on teaching* (pp. 214-229), New York: Macmillan.

Walker, M. (2006), *The choice of Likert or dichotomous items to measure attitudes across culturally distinct countries in international comparative educational research*. Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Walker, M. (2007), Ameliorating Culturally-Based Extreme Response Tendencies To Attitude items, *Journal of Applied Measurement,* No. 8, pp. 267-278.

Warm, T.A. (1989), Weighted Likelihood Estimation of Ability in Item Response Theory, *Psychometrika*, No. 54 (3), pp. 427-450.

Westat (2007), *WesVar® 5.1* Computer software and manual, Rockville, MD: Author (also see *http://www.westat.com/wesvar/*).

Wilson, M. (1994), Comparing Attitude Across Different Cultures: Two Quantitative Approaches to Construct Validity, in M. Wilson (ed.), *Objective measurement II: Theory into practice* (pp. 271-292), Norwood, NJ: Ablex.

Wolter, K.M. (2007), *Introduction to Variance Estimation*. Second edition, Springer: New York.

Wu, M.L., Adams, R.J. & **Wilson, M.R.** (1997), *ConQuest®: Multi-Aspect Test Software* [computer program manual], Camberwell, Vic.: Australian Council for Educational Research.

List of abbreviations – the following abbreviations are used in this report:

ACER	Australian Council for Educational Research
AGFI	Adjusted Goodness-of-Fit Index
BRR	Balanced Repeated Replication
CBAS	Computer Based Assessment of Science
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CITO	National Institute for Educational Measurement, The Netherlands
CIVED	Civic Education Study
DIF	Differential Item Functioning
ENR	Enrolment of 15-year-olds
ESCS	PISA Index of Economic, Social and Cultural Status
ETS	Educational Testing Service
IAEP	International Assessment of Educational Progress
I	Sampling Interval
ICR	Inter-Country Coder Reliability Study
ICT	Information Communication Technology
IEA	International Association for the Evaluation of Educational Achievement
INES	OECD Indicators of Education Systems
IRT	Item Response Theory
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISEI	International Socio-Economic Index
MENR	Enrolment for moderately small school
MOS	Measure of size
NCQM	National Centre Quality Monitor
NDP	National Desired Population
NEP	National Enrolled Population
NFI	Normed Fit Index
NIER	National Institute for Educational Research, Japan
NNFI	Non-Normed Fit Index
NPM	National Project Manager
OECD	Organisation for Economic Cooperation and Development
PISA	Programme for International Student Assessment
PPS	Probability Proportional to Size
PGB	PISA Governing Board
PQM	PISA Quality Monitor
PSU	Primary Sampling Units
QAS	Questionnaire Adaptations Spreadsheet
RMSEA	Root Mean Square Error of Approximation
RN	Random Number
SC	School Co-ordinator
SE	Standard Error
SD	Standard Deviation
SEM	Structural Equation Modelling
SMEG	Subject Matter Expert Group
SPT	Study Programme Table
TA	Test Administrator
TAG	Technical Advisory Group
TCS	Target Cluster Size
TIMSS	Third International Mathematics and Science Study
TIMSS-R	Third International Mathematics and Science Study – Repeat
VENR	Enrolment for very small schools
WLE	Weighted Likelihood Estimates

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table of contents

FOREWORD **3**

CHAPTER 1 **PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT: AN OVERVIEW** **19**

Participation 21

Features of PISA 22

Managing and implementing PISA 23

Organisation of this report 23

READER'S GUIDE **25**

CHAPTER 2 **TEST DESIGN AND TEST DEVELOPMENT** **27**

Test scope and format 28

Test design 28

Test development centres 29

Development timeline 30

The PISA 2006 scientific literacy framework 30

Test development – cognitive items 31

- Item development process 31
- National item submissions 33
- National review of items 34
- International item review 35
- Preparation of dual (English and French) source versions 35

Test development – attitudinal items 35

Field trial 38

- Field trial selection 38
- Field trial design 39
- Despatch of field trial instruments 40
- Field trial coder training 40
- Field trial coder queries 40
- Field trial outcomes 41
- National review of field trial items 42

Main study 42

- Main study science items 42
- Main study reading items 44
- Main study mathematics items 45
- Despatch of main study instruments 46
- Main study coder training 46
- Main study coder query service 46
- Review of main study item analyses 47

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

CHAPTER 3 **THE DEVELOPMENT OF THE PISA CONTEXT QUESTIONNAIRES** **49**

Overview 50

The conceptual structure 51
- A conceptual framework for PISA 2006 51

Research areas in PISA 2006 55

The development of the context questionnaires 57

The coverage of the questionnaire material 58
- Student questionnaire 58
- School questionnaire 59
- International options 59
- National questionnaire material 60

The implementation of the context questionnaires 60

CHAPTER 4 **SAMPLE DESIGN** **63**

Target population and overview of the sampling design 64

Population coverage, and school and student participation rate standards 65
- Coverage of the PISA international target population 65
- Accuracy and precision 66
- School response rates 66
- Student response rates 68

Main study school sample 68
- Definition of the national target population 68
- The sampling frame 69
- Stratification 70
- Assigning a measure of size to each school 74
- School sample selection 74
- PISA and TIMSS or PIRLS overlap control 76
- Student samples 82

CHAPTER 5 **TRANSLATION AND CULTURAL APPROPRIATENESS OF THE TEST AND SURVEY MATERIAL** **85**

Introduction 86

Development of source versions 86

Double translation from two source languages 87

PISA translation and adaptation guidelines 88

Translation training session 89

Testing languages and translation/adaptation procedures 89

International verification of the national versions 91
- VegaSuite 93
- Documentation 93
- Verification of test units 93
- Verification of the booklet shell 94
- Final optical check 94
- Verification of questionnaires and manuals 94
- Final check of coding guides 95
- Verification outcomes 95

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Translation and verification outcomes – national version quality 96
- Analyses at the country level 96
- Analyses at the item level 103
- Summary of items lost at the national level, due to translation, printing or layout errors 104

CHAPTER 6 **FIELD OPERATIONS** **105**

Overview of roles and responsibilities 106
- National project managers 106
- School coordinators 107
- Test administrators 107
- School associates 108

The selection of the school sample 108

Preparation of test booklets, questionnaires and manuals 108

The selection of the student sample 109

Packaging and shipping materials 110

Receipt of materials at the national centre after testing 110

Coding of the tests and questionnaires 111
- Preparing for coding 111
- Logistics prior to coding 113
- Single coding design 115
- Multiple coding 117
- Managing the process coding 118
- Cross-national coding 120
- Questionnaire coding 120

Data entry, data checking and file submission 120
- Data entry 120
- Data checking 120
- Data submission 121
- After data were submitted 121

The main study review 121

CHAPTER 7 **QUALITY ASSURANCE** **123**

PISA quality control 124
- Comprehensive operational manuals 124
- National level implementation planning document 124

PISA quality monitoring 124
- Field trial and main study review 124
- Final optical check 126
- National centre quality monitor (NCQM) visits 126
- PISA quality monitor (PQM) visits 126
- Test administration 127
- Delivery 128

CHAPTER 8 **SURVEY WEIGHTING AND THE CALCULATION OF SAMPLING VARIANCE** **129**

Survey weighting 130

The school base weight 131
- The school weight trimming factor 132

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

- The student base weight 132
- School non-response adjustment 132
- Grade non-response adjustment 134
- Student non-response adjustment 135
- Trimming student weights 136
- Comparing the PISA 2006 student non-response adjustment strategy with the strategy used for PISA 2003 136
- The comparison 138

Calculating sampling variance 139
- The balanced repeated replication variance estimator 139
- Reflecting weighting adjustments 141
- Formation of variance strata 141
- Countries where all students were selected for PISA 141

CHAPTER 9 **SCALING PISA COGNITIVE DATA** **143**

The mixed coefficients multinomial logit model 144
- The population model 145
- Combined model 146

Application to PISA 146
- National calibrations 146
- National reports 147
- International calibration 153
- Student score generation 153

Booklet effects 155

Analysis of data with plausible values 156

Developing common scales for the purposes of trends 157
- Linking PISA 2003 and PISA 2006 for reading and mathematics 158
- Uncertainty in the link 158

CHAPTER 10 **DATA MANAGEMENT PROCEDURES** **163**

Introduction 164

KeyQuest 167

Data management at the national centre 167
- National modifications to the database 167
- Student sampling with *KeyQuest* 167
- Data entry quality control 167

Data cleaning at ACER 171
- Recoding of national adaptations 171
- Data cleaning organisation 171
- Cleaning reports 171
- General recodings 171

Final review of the data 172
- Review of the test and questionnaire data 172
- Review of the sampling data 172

Next steps in preparing the international database 172

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

CHAPTER 11 **SAMPLING OUTCOMES** **175**

Design effects and effective sample sizes 187
- Variability of the design effect 191
- Design effects in PISA for performance variables 191

Summary analyses of the design effect 203
- Countries with outlying standard errors 205

CHAPTER 12 **SCALING OUTCOMES** **207**

International characteristics of the item pool 208
- Test targeting 208
- Test reliability 208
- Domain inter-correlations 208
- Science scales 215

Scaling outcomes 216
- National item deletions 216
- International scaling 219
- Generating student scale scores 219

Test length analysis 219

Booklet effects 221
- Overview of the PISA cognitive reporting scales 232
- PISA overall literacy scales 234
- PISA literacy scales 234
- Special purpose scales 234

Observations concerning the construction of the PISA overall literacy scales 235
- Framework development 235
- Testing time and item characteristics 236
- Characteristics of each of the links 237

Transforming the plausible values to PISA scales 246
- Reading 246
- Mathematics 246
- Science 246
- Attitudinal scales 247

Link error 247

CHAPTER 13 **CODING AND MARKER RELIABILITY STUDIES** **249**

Homogeneity analyses 251

Multiple marking study outcomes (variance components) 254
- Generalisability coefficients 254

International coding review 261
- Background to changed procedures for PISA 2006 261
- ICR procedures 261
- Outcomes 264
- Cautions 270

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

CHAPTER 14 **DATA ADJUDICATION** **271**

Introduction 272
- Implementing the standards – quality assurance 272
- Information available for adjudication 273
- Data adjudication process 273

General outcomes 274
- Overview of response rate issues 274
- Detailed country comments 275

CHAPTER 15 **PROFICIENCY SCALE CONSTRUCTION** **283**

Introduction 284

Development of the described scales 285
- Stage 1: Identifying possible scales 285
- Stage 2: Assigning items to scales 286
- Stage 3: Skills audit 286
- Stage 4: Analysing field trial data 286
- Stage 5: Defining the dimensions 287
- Stage 6: Revising and refining with main study data 287
- Stage 7: Validating 287

Defining proficiency levels 287

Reporting the results for PISA science 290
- Building an item map 290
- Levels of scientific literacy 292
- Interpreting the scientific literacy levels 299

CHAPTER 16 **SCALING PROCEDURES AND CONSTRUCT VALIDATION OF CONTEXT QUESTIONNAIRE DATA** **303**

Overview 304

Simple questionnaire indices 304
- Student questionnaire indices 304
- School questionnaire indices 307
- Parent questionnaire indices 309

Scaling methodology and construct validation 310
- Scaling procedures 310
- Construct validation 312
- Describing questionnaire scale indices 314

Questionnaire scale indices 315
- Student scale indices 315
- School questionnaire scale indices 340
- Parent questionnaire scale indices 342
- The PISA index of economic, social and cultural status (ESCS) 346

CHAPTER 17 **VALIDATION OF THE EMBEDDED ATTITUDINAL SCALES** **351**

Introduction 352

International scalability 353
- Analysis of item dimensionality with exploratory and confirmatory factor analysis 353
- Fit to item response model 353

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

- Reliability 355
- Differential item functioning 355
- Summary of scalability 357

Relationship and comparisons with other variables 357
- Within-country student level correlations with achievement and selected background variables 358
- Relationships between embedded scales and questionnaire 360
- Country level correlations with achievement and selected background variables 361
- Variance decomposition 363
- Observations from other cross-national data collections 363
- Summary of relations with other variables 364

Conclusion 364

CHAPTER 18 **INTERNATIONAL DATABASE** **367**

Files in the database 368
- Student files 368
- School file 370
- Parent file 370

Records in the database 371
- Records included in the database 371
- Records excluded from the database 371

Representing missing data 371

How are students and schools identified? 372

Further information 373

REFERENCES **375**

APPENDICES **379**

Appendix 1 PISA 2006 main study item pool characteristics 380

Appendix 2 Contrast coding used in conditioning 389

Appendix 3 Design effect tables 399

Appendix 4 Changes to core questionnaire items from 2003 to 2006 405

Appendix 5 Mapping of ISCED to years 411

Appendix 6 National household possession items 412

Appendix 7 Exploratory and confirmatory factor analyses for the embedded items 414

Appendix 8 PISA consortium, staff and consultants 416

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

LIST OF BOXES

Box 1.1 Core features of PISA 2006 22

LIST OF FIGURES

Figure 2.1 Main study Interest in Science item 36
Figure 2.2 Main study Support for Scientific Enquiry item 36
Figure 2.3 Field trial Match-the-opinion Responsibility item 37

Figure 3.1 Conceptual grid of variable types 52
Figure 3.2 The two-dimensional conceptual matrix with examples of variables collected or available from other sources 54

Figure 4.1 School response rate standard 67

Figure 6.1 Design for the single coding of science and mathematics 115
Figure 6.2 Design for the single coding of reading 116

Figure 9.1 Example of item statistics in Report 1 148
Figure 9.2 Example of item statistics in Report 2 149
Figure 9.3 Example of item statistics shown in Graph B 150
Figure 9.4 Example of item statistics shown in Graph C 151
Figure 9.5 Example of item statistics shown in Table D 151
Figure 9.6 Example of summary of dodgy items for a country in Report 3a 152
Figure 9.7 Example of summary of dodgy items in Report 3b 152

Figure 10.1 Data management in relation to other parts of PISA 164
Figure 10.2 Major data management stages in PISA 166
Figure 10.3 Validity reports - general hierarchy 170

Figure 11.1 Standard error on a mean estimate depending on the intraclass correlation 188
Figure 11.2 Relationship between the standard error for the science performance mean and the intraclass correlation within explicit strata (PISA 2006) 205

Figure 12.1 Item plot for mathematics items 210
Figure 12.2 Item plot for reading items 211
Figure 12.3 Item plot for science items 212
Figure 12.4 Item plot for interest items 213
Figure 12.5 Item plot for support items 214
Figure 12.6 Scatter plot of per cent correct for reading link items in PISA 2000 and PISA 2003 238
Figure 12.7 Scatter plot of per cent correct for reading link items in PISA 2003 and PISA 2006 240
Figure 12.8 Scatter plot of per cent correct for mathematics link items in PISA 2003 and PISA 2006 242
Figure 12.9 Scatter plot of per cent correct for science link items in PISA 2000 and PISA 2003 244
Figure 12.10 Scatter plot of per cent correct for science link items in PISA 2003 and PISA 2006 245

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Figure 13.1 Variability of the homogeneity indices for science items in field trial 250
Figure 13.2 Average of the homogeneity indices for science items in field trial and main study 251
Figure 13.3 Variability of the homogeneity indices for each science item in the main study 252
Figure 13.4 Variability of the homogeneity indices for each reading item in the main study 252
Figure 13.5 Variability of the homogeneity indices for each mathematics item 252
Figure 13.6 Variability of the homogeneity indices for the participating countries in the main study 253
Figure 13.7 Example of ICR report (reading) 269

Figure 14.1 Attained school response rates 274

Figure 15.1 The relationship between items and students on a proficiency scale 285
Figure 15.2 What it means to be at a level 289
Figure 15.3 A map for selected science items 291
Figure 15.4 Summary descriptions of the six proficiency levels on the science scale 294
Figure 15.5 Summary descriptions of six proficiency levels in *identifying scientific issues* 295
Figure 15.6 Summary descriptions of six proficiency levels in *explaining phenomena scientifically* 297
Figure 15.7 Summary descriptions of six proficiency levels in *using scientific evidence* 300

Figure 16.1 Summed category probabilities for fictitious item 314
Figure 16.2 Fictitious example of an item map 315
Figure 16.3 Scatterplot of country means for ESCS 2003 and ESCS 2006 347

Figure 17.1 Distribution of item fit mean square statistics for embedded attitude items 354
Figure 17.2 An example of the ESC plot for item S408RNA 356
Figure 17.3 Scatterplot of mean mathematics interest against mean mathematics for PISA 2003 363

LIST OF TABLES

Table 1.1 PISA 2006 participants 21

Table 2.1 Cluster rotation design used to form test booklets for PISA 2006 29
Table 2.2 Test development timeline for PISA 2006 30
Table 2.3 Science field trial all items 39
Table 2.4 Allocation of item clusters to test booklets for field trial 39
Table 2.5 Science main study items (item format by competency) 43
Table 2.6 Science main study items (item format by knowledge type) 44
Table 2.7 Science main study items (knowledge category by competency) 44
Table 2.8 Reading main study items (item format by aspect) 44
Table 2.9 Reading main study items (item format by text format) 45
Table 2.10 Reading main study items (text type by aspect) 45
Table 2.11 Mathematics main study items (item format by competency cluster) 45
Table 2.12 Mathematics main study items (item format by content category) 46
Table 2.13 Mathematics main study items (content category by competency cluster) 46

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 3.1 Themes and constructs/variables in PISA 2006 56

Table 4.1 Stratification variables 71
Table 4.2 Schedule of school sampling activities 78

Table 5.1 Countries sharing a common version with national adaptations 90
Table 5.2 PISA 2006 translation/adaptation procedures 91
Table 5.3 Mean deviation and root mean squared error of the item by country interactions for each version 97
Table 5.4 Correlation between national item parameter estimates for Arabic versions 99
Table 5.5 Correlation between national item parameter estimates for Chinese versions 99
Table 5.6 Correlation between national item parameter estimates for Dutch versions 99
Table 5.7 Correlation between national item parameter estimates for English versions 99
Table 5.8 Correlation between national item parameter estimates for French versions 99
Table 5.9 Correlation between national item parameter estimates for German versions 100
Table 5.10 Correlation between national item parameter estimates for Hungarian versions 100
Table 5.11 Correlation between national item parameter estimates for Italian versions 100
Table 5.12 Correlation between national item parameter estimates for Portuguese versions 100
Table 5.13 Correlation between national item parameter estimates for Russian versions 100
Table 5.14 Correlation between national item parameter estimates for Spanish versions 100
Table 5.15 Correlation between national item parameter estimates for Swedish versions 100
Table 5.16 Correlation between national item parameter estimates within countries 101
Table 5.17 Variance estimate 102
Table 5.18 Variance estimates 103

Table 6.1 Design for the multiple coding of science and mathematics 118
Table 6.2 Design for the multiple coding of reading 118

Table 8.1 Non-response classes 133

Table 9.1 Deviation contrast coding scheme 154

Table 10.1 Double entry discrepancies per country: field trial data 169

Table 11.1 Sampling and coverage rates 178
Table 11.2 School response rates before replacement 182
Table 11.3 School response rates after replacement 184
Table 11.4 Student response rates after replacement 185
Table 11.5 Standard errors for the PISA 2006 combined science scale 189
Table 11.6 Design effect 1 by country, by domain and cycle 193
Table 11.7 Effective sample size 1 by country, by domain and cycle 194
Table 11.8 Design effect 2 by country, by domain and cycle 195
Table 11.9 Effective sample size 2 by country, by domain and cycle 196
Table 11.10 Design effect 3 by country, by domain and by cycle 197

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 11.11 Effective sample size 3 by country, by domain and cycle 198
Table 11.12 Design effect 4 by country, by domain and cycle 199
Table 11.13 Effective sample size 4 by country, by domain and cycle 200
Table 11.14 Design effect 5 by country, by domain and cycle 201
Table 11.15 Effective sample size 5 by country, by domain and cycle 202
Table 11.16 Median of the design effect 3 per cycle and per domain across the 35 countries that participated in every cycle 203
Table 11.17 Median of the standard errors of the student performance mean estimate for each domain and PISA cycle for the 35 countries that participated in every cycle 203
Table 11.18 Median of the number of participating schools for each domain and PISA cycle for the 35 countries that participated in every cycle 204
Table 11.19 Median of the school variance estimate for each domain and PISA cycle for the 35 countries that participated in every cycle 204
Table 11.20 Median of the intraclass correlation for each domain and PISA cycle for the 35 countries that participated in every cycle 204
Table 11.21 Median of the within explicit strata intraclass correlation for each domain and PISA cycle for the 35 countries that participated in every cycle 205
Table 11.22 Median of the percentages of school variances explained by explicit stratification variables, for each domain and PISA cycle for the 35 countries that participated in every cycle 205

Table 12.1 Number of sampled student by country and booklet 209
Table 12.2 Reliabilities of each of the four overall scales when scaled separately 215
Table 12.3 Latent correlation between the five domains 215
Table 12.4 Latent correlation between science scales 215
Table 12.5 Items deleted at the national level 216
Table 12.6 Final reliability of the PISA scales 216
Table 12.7 National reliabilities for the main domains 217
Table 12.8 National reliabilities for the science subscales 218
Table 12.9 Average number of not-reached items and missing items by booklet 219
Table 12.10 Average number of not-reached items and missing items by country 220
Table 12.11 Distribution of not-reached items by booklet 221
Table 12.12 Estimated booklet effects on the PISA scale 221
Table 12.13 Estimated booklet effects in logits 221
Table 12.14 Variance in mathematics booklet means 222
Table 12.15 Variance in reading booklet means 224
Table 12.16 Variance in science booklet means 226
Table 12.17 Variance in interest booklet means 228
Table 12.18 Variance in support booklet means 230
Table 12.19 Summary of PISA cognitive reporting scales 233
Table 12.20 Linkage types among PISA domains 2000-2006 235
Table 12.21 Number of unique item minutes for each domain for each PISA assessments 237
Table 12.22 Numbers of link items between successive PISA assessments 237
Table 12.23 Per cent correct for reading link items in PISA 2000 and PISA 2003 238
Table 12.24 Per cent correct for reading link items in PISA 2003 and PISA 2006 239
Table 12.25 Per cent correct for mathematics link items in PISA 2003 and PISA 2006 241

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 12.26 Per cent correct for science link items in PISA 2000 and PISA 2003 243
Table 12.27 Per cent correct for science link items in PISA 2003 and PISA 2006 245
Table 12.28 Link error estimates 247

Table 13.1 Variance components for mathematics 255
Table 13.2 Variance components for science 256
Table 13.3 Variance components for reading 257
Table 13.4 Generalisability estimates for mathematics 258
Table 13.5 Generalisability estimates for science 259
Table 13.6 Generalisability estimates for reading 260
Table 13.7 Examples of flagged cases 263
Table 13.8 Count of analysis groups showing potential bias, by domain 264
Table 13.9 Comparison of codes assigned by verifier and adjudicator 265
Table 13.10 Outcomes of ICR analysis part 1 265
Table 13.11 ICR outcomes by country and domain 266

Table 15.1 Scientific literacy performance band definitions on the PISA scale 293

Table 16.1 ISCO major group white-collar/blue-collar classification 306
Table 16.2 ISCO occupation categories classified as science-related occupations 307
Table 16.3 OECD means and standard deviations of WL estimates 311
Table 16.4 Median, minimum and maximum percentages of between-school variance for student-level indices across countries 313
Table 16.5 Household possessions and home background indices 316
Table 16.6 Scale reliabilities for home possession indices in OECD countries 317
Table 16.7 Scale reliabilities for home possession indices in partner countries/economies 318
Table 16.8 Item parameters for interest in science learning (INTSCIE) 318
Table 16.9 Item parameters for enjoyment of science (JOYSCIE) 319
Table 16.10 Model fit and estimated latent correlations for interest in and enjoyment of science learning 319
Table 16.11 Scale reliabilities for interest in and enjoyment of science learning 320
Table 16.12 Item parameters for instrumental motivation to learn science (INSTSCIE) 320
Table 16.13 Item parameters for future-oriented science motivation (SCIEFUT) 321
Table 16.14 Model fit and estimated latent correlations for motivation to learn science 321
Table 16.15 Scale reliabilities for instrumental and future-oriented science motivation 322
Table 16.16 Item parameters for science self-efficacy (SCIEEFF) 322
Table 16.17 Item parameters for science self-concept (SCSCIE) 323
Table 16.18 Model fit and estimated latent correlations for science self-efficacy and science self-concept 323
Table 16.19 Scale reliabilities for science self-efficacy and science self-concept 324
Table 16.20 Item parameters for general value of science (GENSCIE) 324
Table 16.21 Item parameters for personal value of science (PERSCIE) 325
Table 16.22 Model fit and estimated latent correlations for general and personal value of science 325
Table 16.23 Scale reliabilities for general and personal value of science 326
Table 16.24 Item parameters for science activities (SCIEACT) 326

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.25 Scale reliabilities for the science activities index 327
Table 16.26 Item parameters for awareness of environmental issues (ENVAWARE) 327
Table 16.27 Item parameters for perception of environmental issues (ENVPERC) 328
Table 16.28 Item parameters for environmental optimism (ENVOPT) 328
Table 16.29 Item parameters for responsibility for sustainable development (RESPDEV) 328
Table 16.30 Model fit environment-related constructs 329
Table 16.31 Estimated latent correlations for environment-related constructs 329
Table 16.32 Scale reliabilities for environment-related scales in OECD countries 330
Table 16.33 Scale reliabilities for environment-related scales in non-OECD countries 330
Table 16.34 Item parameters for school preparation for science career (CARPREP) 331
Table 16.35 Item parameters for student information on science careers (CARINFO) 331
Table 16.36 Model fit and estimated latent correlations for science career preparation indices 332
Table 16.37 Scale reliabilities for science career preparation indices 332
Table 16.38 Item parameters for science teaching: interaction (SCINTACT) 333
Table 16.39 Item parameters for science teaching: hands-on activities (SCHANDS) 333
Table 16.40 Item parameters for science teaching: student investigations (SCINVEST) 333
Table 16.41 Item parameters for science teaching: focus on models or applications (SCAPPLY) 334
Table 16.42 Model fit for CFA with science teaching and learning 334
Table 16.43 Estimated latent correlations for constructs related to science teaching and learning 335
Table 16.44 Scale reliabilities for scales to science teaching and learning in OECD countries 336
Table 16.45 Scale reliabilities for scales to science teaching and learning in partner countries/economies 336
Table 16.46 Item parameters for ICT Internet/entertainment use (INTUSE) 337
Table 16.47 Item parameters for ICT program/software use (PRGUSE) 337
Table 16.48 Item parameters for ICT self-confidence in Internet tasks (INTCONF) 337
Table 16.49 Item parameters for ICT self-confidence in high-level ICT tasks (HIGHCONF) 338
Table 16.50 Model fit for CFA with ICT familiarity items 338
Table 16.51 Estimated latent correlations for constructs related to ICT familiarity 339
Table 16.52 Scale reliabilities for ICT familiarity scales 339
Table 16.53 Item parameters for teacher shortage (TCSHORT) 340
Table 16.54 Item parameters for quality of educational resources (SCMATEDU) 340
Table 16.55 Item parameters for school activities to promote the learning of science (SCIPROM) 341
Table 16.56 Item parameters for school activities for learning environmental topics (ENVLEARN) 341
Table 16.57 Scale reliabilities for school-level scales in OECD countries 341
Table 16.58 Scale reliabilities for environment-related scales in partner countries/economies 342
Table 16.59 Item parameters for science activities at age 10 (PQSCIACT) 343
Table 16.60 Item parameters for parent's perception of school quality (PQSCHOOL) 343
Table 16.61 Item parameters for parent's views on importance of science (PQSCIMP) 343
Table 16.62 Item parameters for parent's reports on science career motivation (PQSCCAR) 344
Table 16.63 Item parameters for parent's view on general value of science (PQGENSCI) 344
Table 16.64 Item parameters for parent's view on personal value of science (PQPERSCI) 344
Table 16.65 Item parameters for parent's perception of environmental issues (PQENPERC) 345
Table 16.66 Item parameters for parent's environmental optimism (PQENVOPT) 345

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.67 Scale reliabilities for parent questionnaire scales 345
Table 16.68 Factor loadings and internal consistency of ESCS 2006 in OECD countries 347
Table 16.69 Factor loadings and internal consistency of ESCS 2006 in partner countries/economies 348

Table 17.1 Student-level latent correlations between mathematics, reading, science, embedded interest and embedded support 354
Table 17.2 Summary of the IRT scaling results across countries 355
Table 17.3 Gender DIF table for embedded attitude items 357
Table 17.4 Correlation amongst attitudinal scales, performance scales and HISEI 358
Table 17.5 Correlations for science scale 359
Table 17.6 Loadings of the achievement, interest and support variables on three varimax rotated components 360
Table 17.7 Correlation between embedded attitude scales and questionnaire attitude scales 361
Table 17.8 Rank order correlation five test domains, questionnaire attitude scales and HISEI 362
Table 17.9 Intra-class correlation (rho) 362

Table A1.1 2006 Main study reading item classification 380
Table A1.2 2006 Main study mathematics item classification 381
Table A1.3 2006 Main study science item classification (cognitive) 383
Table A1.4 2006 Main study science embedded item classification (interest in learning science topics) 387
Table A1.5 2006 Main study science embedded item classification (support for scientific enquiry) 388

Table A2.1 2006 Main study contrast coding used in conditioning for the student questionnaire variables 389
Table A2.2 2006 Main study contrast coding used in conditioning for the ICT questionnaire variables 396
Table A2.3 2006 Main study contrast coding used in conditioning for the parent questionnaire variables and other variables 397

Table A3.1 Standard errors of the student performance mean estimate by country, by domain and cycle 399
Table A3.2 Sample sizes by country and cycle 400
Table A3.3 School variance estimate by country, by domain and cycle 401
Table A3.4 Intraclass correlation by country, by domain and cycle 402
Table A3.5 Within explicit strata intraclass correlation by country, by domain and cycle 403
Table A3.6 Percentages of school variance explained by explicit stratification variables, by domain and cycle 404

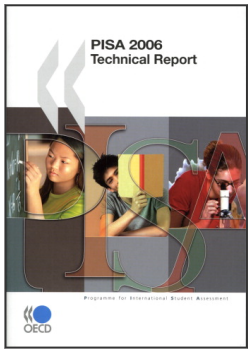
Table A4.1 Student questionnaire 405
Table A4.2 ICT familiarity questionnaire 407
Table A4.3 School questionnaire 408

Table A5.1 Mapping of ISCED to accumulated years of education 411

Table A6.1 National household possession items 412

Table A7.1 Exploratory and confirmatory factor analyses (EFA and CFA) for the embedded items 414

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009



From:
PISA 2006 Technical Report

Access the complete publication at:
https://doi.org/10.1787/9789264048096-en

Please cite this chapter as:

OECD (2009), "Programme for International Student Assessment: An Overview", in *PISA 2006 Technical Report*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264048096-2-en

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.