IERI Monograph Series

Issues and Methodologies in Large-Scale Assessments

SPECIAL ISSUE 2

Measuring Students' Family Background in Large-Scale International Education Studies

October 2013

A joint publication between the International Association for the Evaluation of Educational Achievement (IEA) and Educational Testing Service (ETS)

Copyright © 2013 by Educational Testing Service and International Association for the Evaluation of Educational Achievement.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise without permission in writing from the copyright holder.

ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS) in the USA and other countries worldwide and used under license by the IEA-ETS Research Institute (IERI). IEA and the IEA logo are trademarks and registered trademarks of the International Association for the Evaluation of Educational Achievement and used under license by the IEA-ETS Research Institute (IERI).

ISBN/EAN: 978-088685415-7

Copies of this publication can be obtained from:

IERInstitute
IEA Data Processing and Research Center
Mexikoring 37
22297 Hamburg
Germany

IERInstitute
Educational Testing Service
Mail Stop 13-E
Princeton, NJ 08541
United States

By email: ierinstitute@iea-dpc.de
Free downloads: www.ierinstitute.org

Copyeditors: Paula Wagemaker, Editorial Services, Oturehua, Central Otago, New Zealand, with David Robitaille, Ruth R. Greenwood, Robert Garden, and Sven Lie
Design and production by Becky Bliss Design and Production, Wellington, New Zealand

Editors

Matthias von Davier	Educational Testing Service
Dirk Hastedt	IEA Data Processing and Research Center

Production Editor

Marta Kostek	IEA Data Processing and Research Center

Assistant Production Editor

Chhaya Chavan	IEA Data Processing and Research Center

Reviewers

Daniel Caro	IEA Data Processing and Research Center
Diego Cortés	IEA Data Processing and Research Center
Eugenio Gonzalez	Educational Testing Service
Patrick Gonzales	National Center for Education Statistics
Dirk Hastedt	IEA Data Processing and Research Center
Hans Wagemaker	IEA Secretariat

Foreword

For more than six years, members of the International Association for the Evaluation of Educational Achievement (IEA) and Educational Testing Service (ETS) have worked successfully together within the IEA-ETS Research Institute (IERI) on projects designed to improve the science of large-scale assessments of educational achievement. IERI covers three broad areas of activities: (1) research studies related to the methodology, development, and implementation of large-scale assessments, (2) professional development and training, and (3) dissemination of research findings and information gathered through large-scale assessments.

IERI has published five volumes of the periodical Issues and Methodologies in Large-Scale Assessments, each containing six to seven peer-reviewed papers. The current publication is the second special issue—special because it contains only one (extended) paper. The first special issue addressed a matter highly relevant for researchers planning international large-scale studies, namely the relationship between the sample sizes at each level of a hierarchical model and the precision of the outcome model. Having received very positive feedback on this report issue, we decided to publish another special issue. It addresses a topic also highly relevant to international large-scale assessments—the measurement of students' family background in international large-scale assessments.

Good measures of students' family background are critically important when analyzing large-scale assessment data so as to find which factors are associated with positive outcomes. Various research projects remind us of the high association between family background and students' achievement and also with attitudes and other outcome variables. Researchers therefore try to control for the effect of students' family background when investigating teacher and school effects that may contribute to outcome measures. But finding good measures for family background—especially in an international survey—remains a challenge. We trust you will find this special issue informative and that it help those of you designing studies to develop better measures of family background.

This report is also special because it is the last volume of the IERI monograph series. We decided to convert the series into a SpringerOpen journal called Large-Scale Assessments in Education: A SpringerOpen Journal in order to increase visibility and dissemination of the submitted articles. We hope you will consider submitting papers presenting your own research on international large-scale assessments to our now even more attractive new journal. For further information, please refer to http://www.largescaleassessmentsineducation.com/. Finally, we would like to express our gratitude for the generous support of the research that resulted in this report by the National Center for Education Statistics (NCES) in the United States.

Dirk Hastedt and Matthias von Davier
Editors of the IERI periodical *Issues and Methodologies in Large-Scale Assessments*

ABOUT IEA

The International Association for the Evaluation of Educational Achievement (IEA) is an independent, non-profit, international cooperative of national research institutions and governmental research agencies. Through its comparative research and assessment projects, IEA aims to:

- Provide international benchmarks that may assist policy-makers in identifying the comparative strengths and weaknesses of their education systems
- Provide high-quality data that will increase policy-makers' understanding of key school- and non-school-based factors that influence teaching and learning
- Provide high-quality data that will serve as a resource for identifying areas of concern and action, and for preparing and evaluating educational reforms
- Develop and improve the capacity of educational systems to engage in national strategies for educational monitoring and improvement
- Contribute to development of the worldwide community of researchers in educational evaluation.

Additional information about the IEA is available at www.iea.nl and www.iea-dpc.de.

ABOUT ETS

Educational Testing Service (ETS) is a non-profit institution whose mission is to advance quality and equity in education by providing fair and valid assessments, research, and related services for all people worldwide. In serving individuals, educational institutions and government agencies around the world, ETS customizes solutions to meet the need for teacher professional development products and services, classroom and end-of-course assessments, and research-based teaching and learning tools. Founded in 1947, ETS today develops, administers, and scores more than 24 million tests annually in more than 180 countries, at over 9,000 locations worldwide.

Additional information about ETS is available at www.ets.org.

Measuring Students' Family Background in Large-Scale International Education Studies

Falk Brese and Plamen Mirazchiyski

IEA Data Processing and Research Center

This report was funded by the National Center for Education Statistics under Contract No. ED-08-CO-0117 with the International Association for the Evaluation of Educational Achievement (IEA). Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

IERI Monograph Series
Issues and Methodologies in Large-Scale Assessments

Special Issue 2
Measuring Students' Family Background in Large-Scale International Education Studies

Falk Brese and Plamen Mirazchiyski

TABLE OF CONTENTS

Foreword 3
Acknowledgements 10
Abstract 11
1 Introduction 13
2 Importance of Family Background in Large-Scale Education Studies 17
3 Measuring Family Background 21
3.1 The Development of Measuring Family Background 21
3.1.1 Social class and social stratification 22
3.1.2 Socioeconomic status 23
3.1.3 Cultural and social capital 25
3.1.4 Recent developments in the perception of family background 28
3.2 Indicators of Family Background 28
3.2.1 Income 28
3.2.2 Occupation 30
3.2.3 Household possessions 32
3.2.4 Family structure 33
3.2.5 Immigration status 34
3.2.6 Educational attainment of parents 35
3.2.7 Neighborhood 35
3.2.8 Religion 36
3.3 Indices, Scales, and Other Combinations of Components 37
3.3.1 Socioeconomic status 37
3.3.2 Standard International Occupational Prestige Scale (SIOPS) 38
3.3.3 International Socio-Economic Index of Occupational Status (ISEI) 39
3.3.4 EGP classes 39
3.3.5 Index of Early Home Literacy Activities (EHLA) 40
3.3.6 The PISA Index of Economic, Social, and Cultural Status (ESCS) 41

3.4 The Multilevel Nature of Family Background 41
3.4.1 Individual, classroom, school, and neighborhood 42
3.4.2 Aggregate measures 43
3.5 Administration Issues 45
3.5.1 Reliability of student information about parents 45
3.5.2 Study and questionnaire design 46

4 Research Questions 49

5 Method 51
5.1 Data 51
5.1.1 Studies 51
5.1.2 Countries 63
5.2 Analysis 64
5.2.1 Nonresponse 64
5.2.2 Association with achievement 67
5.2.3 Reliability of scales 68

6 Findings 71
6.1 Home Possessions 73
6.1.1 Response rate 73
6.1.2 Association with achievement 80
6.2 Immigration Status 82
6.2.1 Response rate 82
6.2.2 Association with achievement 83
6.3 Language Used at Home 84
6.3.1 Response rate 84
6.3.2 Association with achievement 85
6.4 Parental Education 86
6.4.1 Response rate 87
6.4.2 Association with achievement 87
6.5 Parental Occupation 88
6.5.1 Response rate 89
6.5.2 Association with achievement 90
6.6 Derived Scales and Indices 90
6.6.1 The PIRLS Index of Home Educational Resources (HER) 91
6.6.2 The PIRLS Index of Early Home Literacy Activities (EHLA) 91
6.6.3 The PISA Home Possessions Scale (HOMEPOSS) 92
6.6.4 The PISA Home Educational Resources Scale (HEDRES) 92
6.6.5 The PISA Cultural Possessions Scale (CULTPOSS) 93
6.6.6 The PISA Index of Wealth Possessions (WEALTH) 93
6.6.7 The PISA Index of Economic, Social, and Cultural Status (ESCS) 94

7 Discussion **95**
7.1 Home Background Domains 96
7.1.1 Immigration status 96
7.1.2 Occupation 97
7.1.3 Parental education 97
7.1.4 Home possessions 98
7.1.5 Scales 98
7.2 Recommendations 99
7.3 Further Research Needs 103
7.3.1 Crosscultural validity 104
7.3.2 Coverage 106
References **109**

Acknowledgements

The authors would like to express their gratitude to numerous people who supported and helped with completing this project. First, we thank the National Center for Education Statistics (NCES), U.S. Department of Education, Institute of Education Sciences, for funding this research project. Furthermore, we are grateful for the support from Hans Wagemaker, Patrick Gonzales, Dirk Hastedt, Leslie Rutkowski, David Rutkowski, Eugenio Gonzalez, Daniel Caro, Diego Cortes, and all reviewers for their valuable feedback. Finally, we would like to thank our colleagues at the IEA Data Processing and Research Center who helped us with the project and also enabled us to spend time on it.

Abstract

The relationship between students' family background and achievement is often seen as an important topic in regard to equality and equity of educational provision. The results of various education studies show that the family background of students correlates with students' academic achievement at school. This paper focuses on the measurement of students' family background within large-scale international studies of education. The intent of this paper is to provide a summary and evaluation of the different ways and concepts of measuring students' family background. It focuses on international large-scale assessment studies using data from the Progress in International Reading Literacy Study (PIRLS) and the Trends in International Mathematics and Science Study (TIMSS) conducted by the International Association for the Evaluation of Educational Achievement (IEA) as well as the Programme for International Student Assessment (PISA) conducted by the Organisation for Economic Co-operation and Development (OECD). In addition to summarizing the different approaches to measuring family background, the paper attempts to evaluate these approaches with regard to the quality of the data collected across countries by using criteria such as response rate, the extent of the relationship between students' family background and achievement, and the reliability of the derived scales. These criteria are also used to compare the different approaches to measuring family background in large-scale international education studies. The paper furthermore tries to identify best practice and provide suggestions on how to improve the measurement of family background within the context of a variety of premises underpinning large-scale international education studies, such as the research questions being explored and the costs of conducting this type of research.

1. Introduction

The results from various education studies show that family background or socioeconomic status (SES) is (sometimes highly) correlated with students' achievement in schools. The influence of family background and family's SES in particular on students' achievement in school has been of considerable interest for a long time. One of the most prominent studies of background characteristics and their relationship with educational outcomes was published by James Coleman and his colleagues in 1966 under the title *Equality of Educational Opportunity*, but it is better known as the Coleman Report (Coleman et al., 1966). This comprehensive exploration of the background characteristics of schools and students that influence the outcomes of education was conducted with a sample of almost 650,000 students and teachers in over 3,000 schools in the United States. The results showed that student background characteristics have a marked impact on the outcomes of education. Although these results were published nearly 50 years ago, social scientists continue to discuss them (see Garoman & Long, 2006).

Family background is often measured using family SES. According to Buchmann (2002, p. 153), three components typically comprise a measure of SES: educational attainment, occupational status, and financial resources. However, as Bourdieu (1986) and Coleman (1988) pointed out, aspects in addition to SES also describe differences between individuals' backgrounds. What these researchers called social and cultural capital are resources that can also reside in the structure and history of the individual's family—hence, family background. Accordingly, in this paper, we do not constrain our analysis to measures of family SES, but extend it to include measures of social and cultural capital. Therefore, family background will be our scope of interest, with family SES included as just one aspect of that background.

In this paper, we attempt not only to identify best practices but also to provide suggestions on how best to measure family background with regard to a variety of premises related to the different content domains, budget and time considerations, target populations, and other features of (future) large-scale international education studies. We furthermore endeavor to give indications about the quality and strength of a scale that is lost when the combination of indicators and components used for analysis is not the best possible one.

More specifically, we explore the different approaches and concepts that large-scale assessments of educational achievement use to measure students' family background, analyze the relationships between family background measures and educational outcomes, and suggest ways to improve large-scale international education study data collection with regard to family background.

Our focus throughout this paper encompasses the measurement of family background within three of the largest international large-scale assessment studies in education: the *Progress in International Reading Literacy Study* (PIRLS), carried out by the International Association for the Evaluation of Educational Achievement (IEA), IEA's *Trends in International Mathematics and Science Study* (TIMSS), and the *Programme for International Student Assessment* (PISA), carried out by the Organisation for Economic Co-operation and Development (OECD). The operationalization of family background varies among the different education studies in scope. For example, while TIMSS mostly reports on single indicators, PIRLS and PISA derive scales (e.g., the PIRLS Index of Home Educational Resources and the PISA Index of Economic, Social, and Cultural Status).

To aid evaluation of the approaches used in these three assessments, we spend the first part of the paper reviewing the development and current use of concepts and indicators of family background, highlighting in particular the different aspects involved in measuring family background. We begin by overviewing developments in the measurement of family background to date, and follow this with an account of the most frequently and prominently used indicators of family background. Next, we introduce several combinations of those indicators that are widely used in the educational research arena. To complete our review of the diverse aspects of measuring family background, we sketch in two additional areas of relevance. These are the multilevel nature of family background and issues associated with administering international large-scale education studies. Our review covers matters additional to those within the ambit of the three international large-scale student assessments. Our purpose at this juncture is to aid identification of possible gaps in or favorable additions to the measurement of family background in those assessments.

During the second part of this paper, we use several criteria to evaluate the different approaches that the international large-scale education studies PIRLS 2006, PISA 2006, and TIMSS 2007 used to measure family background. In order to identify the best (i.e., most desirable) practice (see May, 2002, p. 126), we provide estimates of and numbers pertaining to the quality of the indicators used to measure family background in terms of missing data, reliability, and relationship with achievement. We also, in this regard, consider ability to explain variance in achievement that is lost because of the use of less ideal (sets of) indicators, due to, for example, budget constraints or lack of variables in secondary analysis. We furthermore look at differences across the countries participating in these studies with the aim of identifying possible issues pertaining to crosscultural validity and of assessing the extent to which these measures provide comparable information.

On the basis of this set of analyses, we provide (a) an evaluation of the quality of the indicators of family background used in the above-mentioned international large-scale education studies, and (b) recommendations on the indicators of family background that we consider are the most appropriate for use in future large-scale international education studies and offer the most powerful explanatory power. We end by identifying areas for further research.

2. Importance of Family Background in Large-Scale Education Studies

Policymakers and researchers who invest in and carry out large-scale education studies seek to gain knowledge about core factors associated with differences in students' performance in school. One of the goals of international large-scale studies in education is to provide adequate measures of the background characteristics potentially influencing the outcomes of the education process and to analyze that influence across the countries and subunits of countries[1] participating in these studies.

Given that not all of a child's learning takes place in the classroom, the time a student spends outside school also has the potential to be relevant to his or her academic performance. For example, reading practice and comprehension is not limited to the classroom, as students are likely to encounter texts of various sorts throughout the day, such as advertisements and personal reading materials. Researchers trying to explain students' academic performance need to extend their focus beyond the school to a consideration of the activities students engage in outside of school (see, for example, Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001, p. 21ff.).

A substantial amount of time outside school is shaped by the student's home, that is, the place where the student lives and which includes the people living there. In most cases, children grow up with their parents or guardians, who often influence their children's opinions and views of education as well as their opportunities to learn. In terms of education, the child's family needs to be included as a factor in such child-development processes. In general, parents' and the family's views, values, and morals are, to a great extent, passed on to the child, with that transmittal being a deliberate decision by the parents or an unconscious process. Taking into account this sociocultural reproduction process, which has been identified (Bourdieu, 1977; Bowles, 1977) and discussed in the research literature for some time, brings into play

1 The participants in the international large-scale education studies considered in this paper include not only countries but sometimes also other geographical or political entities that are either subunits of countries or might not be recognized (yet) as sovereign countries by all countries in the world. In this paper, we use the term "country" for reasons of simplicity, with the term comprising all these kinds of geographical or political entities.

a factor that influences the outcome of a child's learning process and is additional to factors at the "core" of the education system, such as curriculum, teachers, classroom, and other school-based elements.

In turn, considering and extracting the effects of family background on student performance in school allows investigation of the influence of school factors on students' achievement "net of family background effects" (Buchmann, 2002, p. 151). Taking into account and controlling for the effect of family background leads to a better understanding of the effects that may be attributed solely to schooling and the formal learning process. Having a clear picture of what schools can and cannot influence has the potential to alter expectations of the outcomes of school reforms or policy programs. Policymakers obtain a better view of the extent of and the limits to which policies directed at the school can influence student performance. Essentially, good measures of family background are an important means of distinguishing the effects of family background from those of the formal education process.

For reasons of simplicity, we refer, in this paper, to "family background" as encompassing circumstances and living conditions grouped together in all kinds of combinations and where students are associated with adults, the latter being parents, guardians, or other adults related to the student. The "parents" do not necessarily need to be married or to live together. Single-parent arrangements and guardians are also included in these home-based living arrangements. We also, in this paper, see the family background of a student as including all those persons who are either in direct contact with that child or young person for a major part of his or her life or who, in some other way, permanently influence his or her living conditions (e.g., due to legal requirements or obligations).

As research over several decades has shown, family background itself is an important variable explaining variance in students' academic achievement. White, Reynolds, Thomas, & Gitzlaff (1993), with their meta-analysis of educational research conducted before 1980, and Sirin (2005), who reviewed similar research between 1990 and 2000, underlined the role family background has historically played in educational research. To this day, the concept of family background remains prominent, not only as a controlling variable but also as a research field of its own. For example, the international reports on PIRLS 2006 (Mullis, Martin, Kennedy, & Foy, 2007), PISA 2006 (OECD, 2007), and TIMSS 2007 (Mullis et al., 2008) dedicated either a complete chapter or at least half of a chapter to family background characteristics and their relationships with achievement.

Comparative analysis between and across countries is, not surprisingly, a feature of large-scale international studies of educational achievement. Countries or education systems participating in these studies are compared to one another with regard to their students' achievement. If countries differ in achievement outcomes, one of the reasons might be the differences in their social structure and in students' family backgrounds, and not (only or partly) because of differences in school-system structures and policies or curricula. Including family background as a potential source

of variation minimizes misinterpretation of research findings. This benefit provides an additional rationale for including family background—at least as a controlling variable—in international large-scale student assessments.

Measuring family background in a diverse set of countries with diverse cultural backgrounds, political systems, and social arrangements creates additional challenges. The measures need to be comparable across countries—that is, and most importantly, valid. For example, if income is used as an indicator of family wealth, how does the amount of money a family earns compare across countries? Clearly, simply using currency exchange rates is not sufficient because differences in the purchasing power of currencies are not taken into account. Also, the value of goods typically differs from country to country. In general, all measures of family background need to be closely examined with respect to their validity and comparability at the international level.

Finally, the question of how students' respective family backgrounds influence achievement tends to be seen not only as a highly important one but also as an issue of equality, debates on which have continued for decades. For example, the "achievement gap" discussion in the United States, initiated by the Coleman Report in the 1960s, continues to be, as noted earlier in this paper, a topic that attracts considerable attention (see Lee, 2002, p. 3f). The U.S.'s *No Child Left Behind* Act of 2001 and the ongoing discussions about it also indicate the relevancy of this topic.

All that we have said so far stresses the importance of measuring family background in international large-scale assessments in education. It is our aim, in this paper, to address these important issues and to provide answers to several questions, such as which of the practices used to measure family background yield, for use by researchers, the most adequate, reliable, and rich information. We also want, in this paper, to try to identify measures that have been proven to collect useful information and those that might be important but are missing from the literature on the international large-scale education studies discussed here. Finally, we want to provide suggestions on what can and could be done to improve the quality of measuring family background.

3. Measuring Family Background

When measuring students' family backgrounds, researchers and analysts need to take several aspects into account. Examples are the inclusion of those backgrounds within the broader theoretical framework, as well as diverse operational issues such as question formats and who to interview (i. e., students themselves, their parents, or other people connected to the students).

Conducting international large-scale assessments in education brings to the fore other specific issues. The international scope of these assessments necessitates checking for cross-country validity—that is, determining if the measures really are measuring something that is comparable across the countries participating in the assessments. Assessing students in schools leads one to query if these students are the most knowledgeable persons to provide information about their families or parents, a concern which then poses questions about the validity of the students' answers. In this chapter, we provide an overview of these aspects of measuring students' family background, keeping our focus throughout on international large-scale assessments in education.

It needs to be noted here that most of the research done on the measurement of family background is based on the work of researchers from the United States and Europe. It is therefore possible that indicators fit better in these parts of the world than in others. This consideration is especially pertinent for international large-scale assessments because the participation of countries from diverse regions of the world is likely to raise issues of validity. Any international assessment needs to be aware of this matter and should carefully validate the indicators that are used.

3.1 The Development of Measuring Family Background

The following section presents a short review of the history of conceptualizing and measuring the characteristics of family background. Included are the occurrence of different approaches and aspects, their theoretical underpinnings, and approaches to measurement that have been evolving over the decades. Some of these characteristics are interrelated because ambiguous theoretical borders sometimes cause researchers and authors to use certain terms interchangeably. Theorists can also differ in how they define different constructs, as can researchers as they seek to operationalize and define the variables used to measure the different constructs constituting the broad concept of "family background."

3.1.1 Social Class and Social Stratification

Social stratification is one of sociology's central terms. It is used to describe the society (or community) in which particular individuals, families, or groups are ranked in hierarchical structures that help determine who controls access to socially valued attributes such as power, wealth, and status. The position of the individual in the hierarchy relates to his or her SES (Mueller & Parcel, 1981). Modern societies can be described as internally divided (or stratified) into groups based on certain characteristics of their members. One of the principal bases and, at the same time, outcomes of such stratification is class. Hobsbawm (1972) realized that classes establish themselves when groups of people acquire the consciousness of themselves as classes. The separation into classes comes into play when the classes relate to one another through interaction in society (Bond, 1981). Max Weber saw social class as a quasigroup or aggregation of individuals governed by its own principles and in close relation to the market. This association of class with the market shifts the meaning of class to the economic position of its members. The individuals share common traits and occupy certain positions in society. According to Warner (cited in Bond, 1981), social class comprises people ordered in socially superior and inferior positions and who are thereby ranked and perceived as different by the members of the community as a whole.

It is a widespread opinion that members of a given social class and their children will reproduce the class itself, with that process relevant not only for individuals but sometimes for whole institutions. Thus, within the sphere of education, schools located in working-class areas tend to prepare their students for working-class jobs (McGregor, 1997). Only a few such students are likely to have the ambition and resources to continue their education in universities (Pearce, Down, & Moore, 2008). The relationship between social-class position and academic achievement in schools was particularly observable until at least the middle of the 20th century in England and in some other European countries. Class origin had considerable influence on children's success in applying for "grammar schools," despite protestations that any child who passed the selection examination could gain entry (Bond, 1981).

Regardless of whether social class has a direct effect on student achievement, differentiation based on social classes may influence, among other things, the individual's aspirations for higher levels of education. As Hatcher (1998) pointed out, even when working-class students have the same level of abilities and achievement as children from higher social classes, they tend to be less motivated to continue their education. Furthermore, parents of children from higher social classes generally have higher expectations than parents from lower classes in regard to their children's achievement (Coleman, 1985), a characteristic that influences school choice and often leads to social and ethnic segregation across schools.

Various studies in education have included social stratification measures, most of which use a single variable, typically occupation, income, or education. Others use scales that combine the information associated with different variables such as education, source of income, type of housing, and type of neighborhood. Striker

(cited in Bond, 1981) criticized these studies for using measures that were not universally applicable to all ethnic groups. Striker's consideration of African Americans in the United States revealed a stratification that was much more complex than the concepts and measures generally used in the studies he critiqued. He listed, as an outcome of his work, over 150 variables associated with social stratification.

3.1.2 Socioeconomic Status

Although there is no strong consensus on the conceptual meaning of SES (see Bornstein & Bradley, 2003), sociologists typically use this term to refer to the relative position of an individual or a family within a hierarchical social structure, based on their access to or control over wealth, prestige, and power (Mueller & Parcel, 1981). This concept is traditionally operationalized through measures characterizing parental educational levels, parental occupational prestige, and family wealth (Gottfried, 1985; Hauser, 1994; Mueller & Parcel, 1981). Over time, the term "social-economic" began to be used to distinguish an individual's background from the previously used term "social class." Later, with the development of numerous other scales, social-economic was reduced to "socioeconomic."

However, as Bond pointed out several decades ago (1981), whether researchers used the term socioeconomic status or social class frequently depended on the researchers' origin: British researchers preferred social class while their U.S. colleagues preferred socioeconomic status. Even so, "class" in British research literature refers mainly to occupation, while SES usually includes a combination of occupation, income, and education. Although Bond admitted that both terms are not equivalent and the difference between them is not clear, he used them as equivalents in his article (Bond, 1981, p. 236).

Mueller and Parcel (1981) disagreed with the interchangeable usage of the terms socioeconomic status and social class, arguing that the individual's or family's SES ranks them on a hierarchy based on their access to or control over valued commodities such as wealth, power, and social status (Mueller & Parcel, 1981, p. 14). Striker (1980, p. 93) added to the debate by claiming that the intention underlying SES as a measure for social stratification was to place individuals on a hierarchy in society. But no matter what meaning was ascribed to this construct, SES continued to be used in various research areas as a variable for reporting characteristics of samples, as a control variable for eliminating extraneous sources of variation in other variables, and as both a predictor of outcomes and as an outcome itself.

Until the end of the first half of the 20th century, measures of SES were quite subjective. When measuring and ranking SES, raters generally took the prestige accorded to an occupation into account. Some occupations were thus placed higher than others simply because of that prestige, rather than for the skills they required or the type of work they embodied (May, 2002).

By the 1980s, a general consensus had emerged amongst researchers that no single variable could be deemed appropriate for measuring SES and that SES is a composite variable typically encompassing education, income, and occupation. Examples are the

SES index used by the U.S. Census Bureau as well as some SES indices used in studies conducted by the U.S.'s National Center for Education Statistics (May, 2002, p. 16). During the 1990s, a number of studies clearly showed that despite the moderate correlation between the three variables, they measured different, albeit important, aspects of SES. Parental income, for example, indicated the likelihood of a student having ready access to resources (both social and economic). Parental education was considered to be the most stable variable because (at least in the U.S.) of its high correlation with income. Occupation, in turn, was seen as related to education and income. By the end of the 1980s and on into the 1990s, an additional indicator appeared in several studies—home resources. This measure was not limited solely to home possessions, such as number of books in the home, availability of a computer, and availability of space in which to study, but also included the availability of educational services after school and during vacation time (Sirin, 2005, p. 419).

Although many studies have shown a correlation between SES and academic achievement, a debate concerning the reasons for this association is ongoing. The debate encompasses four major arguments, all of which emerged in the late 1960s. Proponents of the *genetic argument* state that attributes such as talent are genetically inherited and that certain groups have lower status because they are genetically disadvantaged. Although some of the advocates of this theory admit that society contributes to the individual's development, their major point is that genetic predispositions are the crucial factor (Bond, 1981, pp. 242–243).

According to supporters of the *cultural argument*, children from different socioeconomic groups are placed in different cultural environments, which can influence students' learning because of the different communication and social practices inherent in each of those environments. This argument maintains that "lower-class" children are socialized in more context-tied environments, whereas "upper-class" children are socialized to apply their skills and knowledge to unfamiliar contexts in innovative ways. Critics of the cultural argument point out that the terms "culture" and "motivation" are loosely used, with that usage a reflection of the supposition that the terms lack clear meaning and definition.

The third argument is based on differences in *opportunities to learn*. The thinking here is that variation in student achievement is a result of unequal exposure to knowledge. According to this argument, lower-class children are treated inferiorly in education compared to upper-class students. Although some studies in the 1960s showed that family had more influence on academic success than the school, subsequent research in the 1970s showed that schools themselves may inhibit students' performance, especially those students originating from poor families. For example, in 1966 Leacock claimed (as cited in Bond, 1981, p. 244) that teachers tend to demonstrate, whether consciously or unconsciously, a nonsupportive attitude toward children from lower classes.

The fourth argument used to explain variation in student achievement focuses on *unequal treatment at school*. According to this approach, formal schooling will maintain class inequalities as long as class differences are present in society. Thus, a child who comes from a poor family will demonstrate lower performance in school. According to Bowles (1997), society seeks to maintain its social separation of labor and control over production. Formal education, as embodied in schools, is a means of maintaining this division and thereby ensures political and social stability. This mechanism not only gives more distinct shapes to the classes in a society but also solidifies the values and traits of each, such that they become subcultures (Bond, 1981, p. 245).

3.1.3 Cultural and Social Capital

Beyond economic indicators such as family income, family background is also thought to encompass more ambiguous but important "capital," such as cultural and social capital. Bourdieu (1986) is considered to be the father and the main developer of the theory of forms of capital, although the term "social capital" appeared as early as 1920. In Bourdieu's understanding, capital is accumulated labor that can occur in three forms: economic (can be converted into money and institutionalized into property), cultural (can be institutionalized in educational qualifications and converted into economic capital), and social (is made up of social obligations or connections and can be converted into economic capital and institutionalized).

According to Bourdieu (1986), *cultural capital* exists in three interrelated states. The *embodied* state describes the long-lasting dispositions of the mind and the body. Its acquisition depends on class, period, and society and is unconscious. In parent–child relationships, children's unconscious adoption of their parents' dispositions is part of the reproduction effect that many critical theorists stress. Furthermore, because cultural circumstances shape the acquisition of cultural capital in its embodied state, individuals conducting crossnational research need to take differences in cultures and societies into account. In the *objectified* state, cultural capital objectifies itself in material objects that are seen as the media of the culture, such as paintings, writings, and monuments. But, unlike the embodied form, they can be transferred, just like economic capital. These cultural products can be acquired both materially (as economic capital) and symbolically. In the third state, the *institutionalized* state, cultural capital is objectified in the form of academic qualifications. Academic qualifications are certificates of cultural competence that confer guaranteed cultural value on their owners. Academic qualifications provide their holders with a means of comparing themselves with and against others (Bourdieu, 1986).

Social capital is an aggregate of relations (even only potential ones) within a social network (either institutionalized or informal, or simply presumed) and group membership(s). The groups may be ones that are established institutionally and labeled with a name, for example, family name, school name, political party affiliation. These relations are based on material and symbolic exchanges (Bourdieu, 1986).

The central part of Bourdieu's theory focuses on social networks. Network connections are not ones given by default, naturally, or by social means, but are the product of permanent effort (either conscious or unconscious). The reproduction of social capital presupposes expenditure of time, energy, and economic capital, but cannot be effective without investment of specific competence. For Bourdieu, it was clear that economic capital is the root of all other types of capital. The transformation of economic capital, however, does not happen automatically or spontaneously without labor. Instead, it is a time-consuming pursuit that needs attention and care (Bourdieu, 1986).

Another author who is considered a founder of the theory of social capital is Coleman (1988). He argued that social capital is intangible and exists in three forms: the level of trust evidenced by obligations and expectations, information channels, and norms and sanctions aimed at promoting good for all rather than just for the individual. Like Bourdieu, Coleman stressed the importance of networks, especially the connections across generations. He gave as an example parents who know the parents of their children's friends. In this kind of setting, he saw a social structure that facilitates the emergence of norms.

Coleman (1988) challenged Bourdieu for not distinguishing social capital and the resources or means by which it is obtained. Coleman argued that having greater access to social resources because of their availability within one's own network does not necessarily mean that someone without those advantages in his or her social network is less eligible to procure them. Coleman also disagreed with the notion that the dominant class is able to reproduce itself because of its members' ability to secure social capital. According to Coleman (cited in Dika & Singh, 2002), social capital is social control, but trust, norms, and information channels are also characteristics of a society. As Lareau (2001) explains, Coleman stressed the family's responsibility to adopt certain norms so that their children benefit by having better chances in life, whereas Bourdieu maintained that structural constraints and factors such as class, gender, and race are what influence access to institutional resources.

Compared to the other theories/constructs described so far (SES, for example), the concepts of social and cultural capital are relatively new. Their influence on the field of education, as well as the onset of their measurement, came later, at the end of the 1980s. Nevertheless, as Dika and Singh (2002) point out, the measurement of social capital that took place in educational research was comparatively easy and straightforward. In 1988, Coleman used data from the High School and Beyond (HSB) study to explore the relationship between social capital and educational outcomes. He used different measures (both parents at home, number of siblings, parental educational expectations, closeness between the generations) to argue that higher accrual of social capital leads to lower rates of school dropout.

Many subsequent educational studies adopted Bourdieu's approach for measuring social capital. The researchers involved in these studies included, as measures of social capital, language used in classroom instruction, career decisionmaking, academic

discourse, and relations between schools and families. Educational sociologists have also applied Bourdieu's approach to social and cultural capital in their research directed toward explaining differences in schools, and using class, gender, race, and ethnicity as the characteristics of interest (Dika & Singh, 2002). De Graaf, De Graaf, and Kraaykamp (2000) investigated the effects of parental beaux arts participation (i.e., attending art galleries, museums, opera, ballets, plays, cabarets, and classical music concerts) and parental reading behavior (as aspects of parental cultural resources) on educational attainment. Lee and Bowen (2006) used different aspects of parental involvement at school to explore the influence of these aspects on students' achievement. They distinguished levels of parental involvement in terms of differences in parents' habitus and considered the extent to which the effects of parental involvement reflected differences in cultural capital. Although our paper concentrates solely on quantitative approaches, it is pertinent to note that Bourdieu's and Coleman's concepts of economic, social, and cultural capital have also been used in qualitative and ethnographic approaches (see, amongst others, Lane & Taber, 2012; Lareau, 1987; Lareau & Weininger, 2003).

As Bourdieu (1986) himself stated, the best measure of cultural capital is the time invested in obtaining it. The reason why it is the best relates to the fact that time needs to be invested to transform economic capital into cultural capital, which, in turn, again needs economic capital. The cultural capital of a family is thus obtained by investing time, and time can also be spent spreading it among family members—on transferring it from one family member to another (Bourdieu, 1986). However, for Vryonides (2007), measuring both social and cultural capital raises the issue of operationalization of the constructs to be measured and the selection of variables. Vryonides delineates the different understandings of social capital between the two major theorists (Bourdieu and Coleman) in the following way: Bourdieu saw social capital as an aggregate of resources (actual or potential), linked to possessions steadily accumulated within a network of institutionalized relationships, whereas Coleman located social capital within the family (as a structure and as relationships between generations) and outside of it (the social ties outside the family that create trust and obligations and impose norms).

Sullivan (cited in Vryonides, 2007) contradicts Bourdieu's assumption regarding cultural capital that higher parental education indicates higher levels of cultural capital. She states that such an assumption is inadequate and misses the broadness of the concept. Cultural capital, she continues, has been operationalized in many different ways by many different researchers. As Vryonides (2007) points out, the most likely reason for this diversity is that the concept of cultural capital is very broad and not easy to quantify. Some researchers at the end of the 20th century and the beginning of the 21st century used measures such as frequency of visiting "high" cultural events (e.g., classical concerts, opera, museums, and art galleries) as well as reading habits in the family. For Vryonides (2007), this kind of approach is too narrow a way of measuring cultural capital at home, especially with respect to Bourdieu's definition of the concept.

When formulating and conducting their international large-scale assessments, both IEA and the OECD have attempted to operationalize and measure cultural and social capital. For example, the TIMSS 1995 framework divided family background into forms of capital (see Martin & Kelly, 1996, pp. 5–6). The PISA 2000 framework adopted a similar approach by referencing "students and their family backgrounds, including the economic, social and cultural capital of students and their families" (OECD, 1999, p. 15).

3.1.4 Recent Developments in the Perception of Family Background

There seems little doubt that the position of a person or group (e.g., family) in the social hierarchy is an important consideration in any discussion on different opportunities in society. For example, belonging to a certain social class and having achieved a certain occupational status can be indicators of stability in the socioeconomic hierarchy. However, research done by May (2002) shows that family-specific background indicators provide important information about the resources associated with educational outcomes that are available to a family at a particular point in time. Take, for example, the occupational status of one or more family members. Occupational scales—especially those that are highly aggregated—can mask important variations of interest simply by grouping similar occupations into larger categories without taking into account the differences in education or the current economic situation of the people who belong to particular occupational categories. Using measures directly concerned with the individual's family resources can reduce the measurement error associated with aggregated measures. Thus, we can obtain more precise estimates of the family resources that are relevant to educational outcomes. Hence, "... a shift should occur in contemporary SES theory from one emphasizing class distinctions and/or positions in a social and economic hierarchy to one emphasizing individual or family resources at a specific point in time" (May, 2002, p. 131).

3.2 Indicators of Family Background

In this section, we select from the wide array of possible indicators of family background and family characteristics, with our selection focused on those indicators that (a) have been most prominent in recent educational research, and (b) that have proven to be compatible with theoretical approaches to and constructs of family background. We evaluate each indicator with regard to its explanatory power and connectedness to theoretical approaches. We also consider research-based administration issues, especially those associated with international large-scale student assessments. Although not all indicators discussed here have been used in TIMSS, PIRLS, and PISA, we mention them because of their prominent role in educational research in general and their potential utility in international assessments in particular.

3.2.1 Income

Family income is frequently used as an indicator of family wealth and is seen as an important measure of family background. As Perl (1973) pointed out, parents from high-income families are able to provide their children with learning resources, such as books and other educational materials, a place and a time to study, and

a neighborhood of same-income families and children from well-educated families. He argued, on the basis of his empirical research, that students from high-income families are likely to attend schools where the other students come from a similar background, a situation that augments their educational achievement. According to Perl, this opportunity also applies to high-income families where parents are not so well educated (Perl, 1973).

For Hauser (1994), the majority of studies measuring income as part of family background or SES improperly use variables related either to the father or the mother. Such misuse can result in biased information, especially in cases where children do not live in a family with both biological parents (e.g., one biological and one step-parent, single parents, foster-parents). Hauser suggested that in addition to measuring a family's income, researchers should focus on the "householder" (Hauser, 1994, p. 1542), that is, the head of the household who is also the main earner, regardless of gender, and collecting additional information on his or her educational attainment, labor force status, and occupational position. In Hauser's opinion, for most children the reference person would be the father or the father substitute in households where both parents are present.

The family income or, as defined by Duncan and Magnuson (2003), "household income," is the sum of all sources of income that come into a family and that all its members receive over a certain period of time. This reference period is most frequently a month or a calendar year. When combined with wealth measures, household income gives information on the ability to provide children living in the household with their basic needs as well as a healthy and safe environment, which can also have an effect on children's motivation. According to Duncan and Magnuson (2003), household income should be adjusted for sources such as food stamps, tax credits or supports, and paid taxes, an adjustment that would refine the measurement. Duncan and Magnuson claim that the division of income among the members of a household, which they term the "income-per-need-ratio," shows the per-capita disposal of the resources even more clearly.

As with any other indicator, there are some issues with measuring income, such as coverage of the sources of income. How detailed should the information be? What sources of income should be covered? Hauser (1994, p. 1543) concluded that total household income during the year preceding the survey is the most important information and, therefore, sufficient. Alternatively, each individual's earnings and a limited set of questions concerning social supports, such as receiving subsidies for housing or procuring food stamps, could be used.

Compared with other measures of social status or occupation, income is harder to measure because it can vary. It not only changes with wages and other sources of income, but also with the change in the composition of members in the household and in the employment status of particular family members. Furthermore, while other measures are more stable over time, such as parents' education, income is not always stable or predictable for certain segments of society.

Another problem that arises in regard to measuring household income lies with missing data due to nonresponding participants, an absence which also affects the reliability of the measure. Participants in surveys can be very sensitive about questions regarding their economic situation and often refuse to answer when they are asked to provide more precise information. For example, in the National Survey of Family and Households reported by Demo and Acock (cited in Ensminger et al., 2000), 17.5% of the mothers did not report their family income.

Another challenge with income as an indicator lies with comparability across countries participating in international comparative research. The income of families can easily be compared on a national level. But at the international level, absolute income is difficult to compare because of different currencies, let alone fluctuations in currency exchange rates.

In educational research, an additional issue regarding the administration of questions about family income emerges. Children might simply not know about their parents' income. Thus, the reliability of the information provided becomes questionable. This issue could be addressed by surveying the parents directly. However, along with potential problems arising out of respondents' unwillingness to provide information about such sensitive data, there is the additional effort needed to develop and administer a parental questionnaire, not to mention the likely significant rise in administration costs. Overall, there seems to be a good many issues associated with using income as a variable of family background.

Researchers have attempted to address these various problems by using certain indicators as a proxy for income, such as poverty status, housing tenure, or participation in school-based free lunch programs. However, as Hauser (1994) discussed, these proxy measures introduce additional problems that make their utility less than ideal, with crosscountry validity being the most critical one.

3.2.2 Occupation

Another often-used indicator of family background is occupation. The traditional approach when measuring occupation (along with other characteristics) is to use only the occupation of the head of the household. Usually, this person is assumed to be the father or the "father figure"—the main person in the household responsible for the socioeconomic wellbeing of the family. When the father's occupation is not available, the mother's occupation is often taken as a substitute. But the relevance of this approach has become increasingly questionable. The employment situation in households has changed significantly over the last decades, especially with mothers increasingly contributing to family income. Nevertheless, as Entwisle and Astone (1994) have pointed out, even though women in modern society have jobs, and often the job is one with relatively high occupational prestige, they still tend to be paid less than men.

Today, the theory and practice of measuring SES-related components emphasizes the considerable importance of occupational status compared to financial resources and education. The stability of this indicator, as already discussed, is the main reason

for this change. A stable occupation status shows a certain position in the hierarchy of a society and also shows the economic situation. However, May (2002, p. 130) suggests that more variant indicators contain additional important information related to educational outcomes and wealth. Different family indicators may reveal important facts about the current economic situation that change over time, while occupational status scales may mask this important variation. Usually, for analysis purposes, researchers group similar occupations into larger categories that can mask differences in educational attainment or the current economic situation underlying these occupations. Scales that contain family-specific indicators normally reduce the measurement error arising from the aforementioned problem.

Earlier in this paper, we stated that occupation is a preferable measure compared to income because it is more stable over time. But is occupational status really so stable? In 1995, Swinnerton and Wial published their research on job stability among low-seniority workers in the U.S. The data showed that occupational stability in general can decline for certain periods of time depending on the factors related to what the two authors called a "business cycle." The authors suggested that there is a general tendency toward increasing job instability in the U.S. (Swinnerton & Wial, 1995, p. 303). Nevertheless, occupational status still remains one of the most stable indicators of family socioeconomic status.

Occupation has huge potential as a measure of family background. It could be operationalized as an indicator of family wealth or prestige. It could also relate to social capital, given that some occupations imply a higher level of connectedness with other people or societal institutions. That said, obtaining the information about parental occupation has some challenges. While respondents usually do not see their occupation as sensitive information, unlike income, and (older) students can usually provide reliable occupation information with regard to their parents (see Chapter 3.5), occupational data always need to be transformed into a measure that can be used for comparisons.

Usually, occupation data are grouped and afterwards scaled (see Chapter 3.3). Depending on the purpose of the research, this grouping can differ from study to study, reducing comparability. Furthermore, in international settings, the ascription of value, prestige, or other characteristics used for comparison purposes can vary around the world. Using occupation at the international level therefore requires knowledge about these differences. Furthermore, these ascriptions generally change over time. Lastly, the business world is constantly developing, creating new jobs and changing the world of labor, which poses yet another challenge to comparisons across time. Nevertheless, data on occupation are relatively easy to collect, researchers worldwide are addressing the challenges mentioned above (see especially the authors deriving indices and scales from the occupation data mentioned in Chapter 3.3), and information on parental occupation has shown, overall, to be a valuable indicator of family background both in the past and currently.

Industry

Several authors, among them Duncan and Magnuson (2003), Entwisle and Aston (1994), Hauser (1994), and Hauser and Warren (1997), argue that using only occupational status and job type is not sufficient, and that researchers should also collect information about the industry in which each parent works. The issue that is of interest here relates to occupational prestige. As Entwisle and Astone (1994) state, for some occupations, job prestige depends on the industry itself. Researchers must know the industry to properly assign the occupational prestige of parents. Industry, according to Hauser and Warren (1997), is an abstract category. It is used to group and classify products and services, despite there being complex interdependencies between job and industry classifications. Hauser and Warren discuss in detail the necessary coding procedures, which, as they demonstrate, can be complex and time consuming.

The problems related to collecting data on industry are manifold. As Entwisle and Astone (1994) remind us, industry is rarely, if ever, used directly in analyses of data, other than in terms of defining the prestige of an occupation. Furthermore, children cannot be reliably expected to provide the name of the industry or industries their parents work in. Finally, the issue of comparability of industry classifications at the international level remains. Categories of industry are likely to have a different connotation or even a different prestige in different countries.

3.2.3 Household Possessions

When discussing cultural capital, we mentioned that its objectified state concerns the material objects that serve as a media for a culture (e. g., paintings, writings, monuments, etc.). According to Yang (2003), the number of books at home (among other cultural possessions) is hypothesized to be a measure of the cultural and educational resources at home, although these indicators vary greatly across countries with regard to the cultural value they represent. Furthermore, the number of books at home indicates parents' emphasis on education or intellectual activities beyond books used as educational resources. As Elley (1992, p. 14) pointed out, "The availability of books is a key factor in reading literacy. The highest scoring countries typically provide their students with greater access to books in the home, in nearby community libraries and book stores, and in the school."

In large-scale education studies, students are often administered questions concerning their home possessions, household composition, and parents. Yang and Gustafsson (2004) argue that data on home possessions collected from young children are much more reliable compared to the information they provide about their parents' education, jobs, and income. It seems that children find it much easier to grasp the notion of the material possessions they have daily contact with at home than to understand notions of parental education, job, and income. So, regardless of the criticism about the reliability of data on possessions provided by students, we can assume it is a valuable source of information in the absence of parents either not being asked or not answering questions about these matters.

Comparisons of possessions across different countries are, however, problematic. International large-scale education studies generally use various indicators to measure home possessions as part of family background, but these indicators are not universal across all countries. They may be valued differently in any two societies because of cultural diversity and economic differences. If some possessions are valued highly in a country because of limited access to them (for whatever reasons), in other societies the same possessions may not be a good indicator at all because they are widely available and therefore found in most households (Yang & Gustafsson, 2004).

Some international education studies even administer different sets of indicators of home possessions in different countries. The designs of studies such as PIRLS, TIMSS, and PISA provide countries with the opportunity to add to the set of possessions cited in the questionnaires in all countries their own local indicators of home possessions that their cultures value highly. While this approach suits a country's own cultural and economic situation, it can, at the same time, pose problems for international comparisons. Not all countries add the same number of these optional indicators, a situation which ultimately results in different numbers of indicators across countries. Another problem occurs when items are very different in their nature (Yang, 2003). For example, for TIMSS 2007, Dubai added a private driver and a private maid to the list of home possessions. Armenia added a bible as an indicator. These examples not only show how different societies ascribe value to different indicators but also how the items that are valued may depend on the wealth of a society or country.

3.2.4 Family Structure

Family structure is defined as the composition of members of the family. Usually, the people living together with the child in the household is the most interesting characteristic of family structure, as these people will typically have an influence on the child's development and living conditions (see Chapter 2). Family structure is often used synonymously with household composition. Important aspects are the presence or absence of parents living together with the child in the household. Often, the presence of siblings and grandparents is also seen as relevant.

As we stated earlier, the traditional strategy of measuring household income, wealth, or occupation suggests using measures that take into account the "household head," generally perceived to be the male providing the resources to the family. Nowadays, the situation is much more complex because many women have jobs that enable them to contribute much to family income and social status. When we take into account gender differences and the prestige of the jobs in which men and women are involved, the picture becomes even more complex. Additionally, the many single-parent families and other types of family structure (e.g., female-headed families) make researchers doubt the utility of the traditional approach to providing unbiased measures (Mueller & Parcel, 1981).

Nevertheless, family structure can serve well as an indicator of social capital. In general terms, the bigger the family is, the more potential there is for establishing connections and building on relationships with other family members. This characteristic is even

more valid in societies with cultures that are family oriented, such as those in the Arabian region (see, for example, Joseph, 1994). If we regard family size as one aspect of family structure, we again need to take cultural differences into account. A huge family can be seen as advantaging a family member in terms of social connectivity, but also as disadvantaging him or her from an economic point of view, because all family members' living costs need to be taken into account. In developed countries, especially, a large number of children might be regarded as undesirable.

3.2.5 Immigration Status

The world's demographics of today are much more dynamic compared to the period of about 100 years ago. Migration of people, for example, is more frequent within and across countries. This movement across national borders produces many challenges for the education of immigrants and their further development. Immigration status is an important characteristic of students because it is usually related with low family SES and hence the risk of low academic achievement and low educational attainment of students, as well as of school dropout (Kao & Thompson, 2003). Students with an immigrant background may experience difficulties adjusting to the new environment and to the new culture as well. Immigrant children often speak a native language that is different from the language of instruction in the schools they attend, thereby presenting them with an additional burden. As Mullis and her colleagues mention, there are countries in which students are "at double disadvantage due to their parents' education and socioeconomic background" (Mullis et al., 2009, p. 114). International studies in education usually administer questions to students about their own and their parents' immigration status. For example, the TIMSS 2007 student questionnaire asked students about the frequency with which the language of the TIMSS test was used in their homes and whether they and each of their parents were born in the country in which they were now residing. Through questions such as these, a questionnaire can gather information on the language spoken at home and whether a student is a first- or second-generation immigrant.

The student's and family's immigration status can contribute to both economic and social capital. The influence of migration on the economic situation and social connectedness of the family is, to some extent, related to the reason for migrating to another country. For example, refugees are more likely to leave not only their country but also a certain economic standard the family might have achieved and the social ties they have built up over the years. Immigration status can therefore indicate a relatively clear and substantial change in these two characteristics.

The influence of the family's immigration status on the student's achievement may also depend on the country's immigration policies. For example, special mandatory programs can ensure that people learn to speak the official language of the country they move to. Policies directed toward integrating immigrant families can influence social integration by ensuring access to social goods and welfare as well as economic status by ensuring easy access to the labor market. The "success" of such programs (i.e., whether or not the programs yield the desired results) is, however, a different issue.

3.2.6 Educational Attainment of Parents

There is a clear indication from research in education that educational attainment of parents is related to students' academic achievement: The higher the parental attainment is, the higher their children's achievement tends to be. Some studies from the past show that the mother's educational attainment has a stronger relationship with student achievement than does the father's educational attainment (Gorman & Yu, 1990). Additional to the effect on student achievement are the strong correlations between parental educational attainment, parental teaching styles, home-learning environment, and student behavior (Duncan & Magnuson, 2003). Studies on educational achievement usually also seek information about the educational attainment of parents. PIRLS, TIMSS, and PISA, for example, all include questions about the highest level of education attained by each student's parents.

Parents' educational attainment is one of the central characteristics of family background with regard to explaining variation in student achievement. It is an indicator of the institutionalized state of cultural capital, because educational achievement is usually awarded with certificates, which then become prerequisites for access to labor markets.

Collecting information at the international level on parental educational attainment has become a standard practice in comparative educational research. In order to allow valid comparison of the variety of educational attainments and (school) education pathways evident worldwide, UNESCO developed a classification scheme, the International Standard Classification of Education (ISCED), which is well proven, accepted, and commonly used around the world. It was last updated in 1997 (UNESCO, 1997).

3.2.7 Neighborhood

Neighborhood characteristics, such as neighborhood income and SES, as well as the level of social (dis-)organization are also important characteristics of a student's family life. Neighborhood SES is usually measured as the proportion of people of 20 years of age who have, according to their nation's census, not completed high school, but this measure is usually seen as a weak one. Neighborhood SES and family SES tend to be closely related. Family SES, for example, generally indicates the type of area in which the family resides as well as the school that the children will attend. Family, as previously discussed, also directly or indirectly determines the provision of home resources and social capital—the supportive relationships among structural forces and individuals.

Some studies have found that the predictive power of neighborhood characteristics on achievement is stronger than the predictive power of family SES (see, for example, Sirin, 2005). Similarly, schools that have higher percentages of students from high-SES families maintain a better peer culture and have a better learning environment, advantages which often lead to higher academic performance (Yang-Hansen, 2008). Additionally, relationships appear to exist between the characteristics of the

community, the school's SES, and the academic achievement of students, with these relationships being strongest in decentralized education systems (Yang-Hansen, 2008).

The theory of social disorganization states that the structural characteristics of the neighborhood define its social organization (see Kohen, Brooks-Gunn, Leventhal, & Hertzman, 2002). Such characteristics are poverty, residential instability, ethnic heterogeneity, and single parenthood. The social organization of a neighborhood defines the norms and values within it as well as the behavior of the residents and the state of public order. Most empirical studies use mainly census data to elucidate the sociodemographic characteristics of the neighborhoods in which families live. These studies adopt the variables mentioned above and include unemployment. Some researchers have criticized this approach in terms of it gathering incomplete data (see Kohen et al., 2002). The alternative approach focuses on collecting information through community surveys and systematic social observations. Kohen et al. (2002), among other researchers, suggest surveying cities in order to examine neighborhood context. They also promote conducting ethnographic studies in order to examine in depth the social structure of particular neighborhoods. Such ethnographic studies usually focus on low income and urbanization.

The criterion of neighborhood in international large-scale education studies is also conceptualized as an attribute of the school rather than of the individual. In PIRLS and TIMSS, school principals are asked to provide information on the proportional distribution of some family background characteristics among the students attending the school. For example, TIMSS 2007 collected information about the percentages of students coming from economically disadvantaged homes, as perceived by the school principal. On average across the participating countries, this criterion showed a statistically significant negative relationship with mathematics achievement (Mullis et al., 2008, pp. 317–318.). However, because it is an aggregate measure at the school level, it cannot be used as an indicator of family background at the individual level.

3.2.8 Religion

When measuring socioeconomic status from a crossnational perspective, Wolf (2005) included variables concerning religiosity. The relationship between individuals and religious groups can be described in two ways—first, whether the respondent is a member of a religious group and, second, whether the respondent identifies himself or herself with a religious group. The term "membership" needs to be refined, however, and not left to the respondent to interpret its meaning.

Religion was included in IEA's International Civic and Citizenship Education Study (ICCS) 2009 assessment as an indicator for students, not families (see Schulz, Ainley, Fraillon, Kerr, & Losito, 2010). But because religious affiliation is usually transferred from parents to their children and because that transference may not hold for the student's actual religious beliefs, it might better serve as an indication of the family's tradition regarding religious affiliation.

Similar to income, information about religion is sometimes seen as sensitive. Additionally, data protection regulations in several countries do not allow researchers or other agents to collect this kind of information. In ICCS 2009, for example, questions about religion were designed as national options: 28 of the 38 participating countries chose to administer these questions (Schulz et al., 2010, p. 33).

3.3 Indices, Scales, and Other Combinations of Components

In order to collect information on family background, researchers have developed various scales designed to measure various family characteristics. The broadest one in terms of concept is socioeconomic status (SES), which is a more or less overarching name for various and quite different measures, including all possible combinations of social and economic indicators. Much more clearly defined indices and scales exist that include parental occupation. Occupational positions in the stratification system of sociology are measured using one of the following three approaches: prestige ratings, sociologically derived class measures, and socioeconomic status scores. For each of these approaches, measures exist that are widely used in international comparative research. One such is the Standard International Occupational Prestige Scale (SIOPS). It measures occupational prestige according to a rating of occupations. Another measure, the EGP (named after its authors—Erikson, Goldthorpe, and Portocarero), is an occupational class scheme. The International Socio-Economic Index of Occupational Status (ISEI) follows the third approach—a socioeconomic status score (Ganzeboom, De Graaf, & Treiman, 1992).

Two additional types of scales associated with PIRLS and PISA receive more consideration at the end of this section. Suffice to say here that the PIRLS Early Home Literacy Activities (EHLA) scale was developed to measure activities and resources in the child's family relative to the child's reading capability (Martin, Mullis, & Kennedy, 2007, p. 201), while PISA's Index of Economic, Social, and Cultural Status (ESCS) endeavors to capture a combined measure of all three forms of capital (OECD, 2009, p. 346ff.).

3.3.1 Socioeconomic Status

Socioeconomic status (SES) is by far the most prominent and widely used latent construct for measuring family background. It is also the least well-defined concept. Introduced into the U.S. in the late 1960s, SES as a construct was recognized, by the 1980s, as problematic because researchers were underpinning it with a variety of different combinations of variables, creating ambiguity in the interpretation of research results. The same conclusion holds nearly a quarter of a century later (see Sirin, 2005). As we discussed earlier, many researchers continue to use the terms SES and social class interchangeably when referring to the social and economic characteristics of the student (or rather the student's family), but they neither clarify their understanding of these terms nor provide a rationale for using them in this way.

According to Merola (2005), measures of SES are usually derived from educational attainment, occupational status, and financial resources. However, not all studies use all three components. For example, the U.S.'s National Assessment of Educational

Progress (NAEP) studies use only the educational attainment and financial resources information from the data collected from students, not parents, which has always been seen as problematic (Merola, 2005). Problems with using only two of the three components have also been reported by Freidlin and Salvucci (1995).

In contrast, there are SES constructs that include more than the above-mentioned three "core" components. One such encompasses cultural aspects. The economic and cultural aspects of SES do not exist separately within a family, but are related. As Yang (2003) points out, cultural indicators measure more than just the cultural resources in the household: some cultural possessions are, alongside their sociocultural role in the home, relatively expensive and require sufficient economic resources for gathering them. This applies to a lesser extent to books and daily newspapers or magazines, and much more to items such as paintings and musical instruments. Cultural possessions in general play a significant role in social and economic distinctions. Hence, economic and cultural aspects of SES are seen as interrelated (Yang, 2003).

3.3.2 Standard International Occupational Prestige Scale (SIOPS)

The Standard International Occupational Prestige Scale (SIOPS) was developed and published by Donald Treiman (1977). SIOPS is based on 509 occupations, distributed in 11 major groups and then divided into 84 minor groups. The latter, in turn, are divided into 288 units. The occupations with the highest prestige are those at the top of the political, religious, and educational hierarchies. The lowest prestige is assigned to occupations that are illegal or illegitimate (e.g., drug peddler or smuggler). Treiman (1977) reviewed 85 occupational prestige studies conducted in 60 societies all over the world during the 20th century up to 1971. He chose his data sources mainly on the basis of the quality of the samples. The classification of the occupations across countries was based on the International Standard Classification of Occupations (ISCO) developed by the International Labor Office (ILO) in 1968 (ILO, 1968). Due to incomplete matches of the occupations from country to country, Treiman developed a criterion for matching occupations crossnationally. His work took into account job titles, the tasks that particular jobs involve, and the functions within those jobs (Treiman, 1977). Treiman validated the scale, so enabling its users to explore the occupational systems as well as particular occupations on their own, and from there classify individuals according to their respective occupations. During the 1990s, Treiman and his colleague Ganzeboom updated the index, using the revised ISCO-88 (ILO, 1990) (see Ganzeboom & Treiman, 1996).

Prestige scales such as the SIOPS pose problems when being used for international comparison purposes. This is because different societies may value occupations differently, thus ascribing different prestige to the same occupation. The prevailing sector of an economy might influence the prestige of (certain) occupations. For example, in a society where agriculture is the predominant sector, agricultural occupations might be valued (much) more than in developed countries where the service industry prevails.

3.3.3 International Socio-Economic Index of Occupational Status (ISEI)

The International Socio-Economic Index of Occupational Status (ISEI) was developed by Ganzeboom et al. (1992). They wanted to create an internationally comparable socioeconomic index scale that took into account the association between SES and occupational prestige. Prestige scales consist of evaluative judgments. Socioeconomic indices are not, however, based on such judgments but rather are constructed through calculation, for example, by weighting the sum of average education and average income. Ganzeboom and his colleagues used a combination of data collected between 1968 and 1982 from 31 different studies conducted in 16 different countries by the ILO. They derived the index by using occupational titles coded according to the ISCO-68 scale and also including education and income, and restricted their data to those pertaining to male respondents only, 21 to 64 years of age, and working fulltime (i.e., 30 or more hours per week).

The three colleagues then derived occupational groups from the data using the ISCO-68 categories to define occupational units. They found the education stratification measures difficult to develop because of the diversity of education systems across countries, but eventually derived their measures from two different sources—the number of years spent in formal study and the type of education completed. The researchers also encountered problems with respect to comparability of income measures, again because of crossnational differences. However, they decided to divide incomes in the datasets by their means and apply a logarithmic transformation, after which they transformed the result into a Z-scale with a mean of zero and a standard deviation of one (Ganzeboom et al., 1992). This index also received a major revision after the ISCO-88 classification was made available (see Ganzeboom & Treiman, 1996).

Individuals assigning ISEI values need only the information about occupation. After coding the occupation to the ISCO-88 classification, all they need to do is conduct simple recodings that transform the ISCO-88 codes into ISEI values, thus making the ISEI scale very easy to use. It also enables transfer of nominal data on occupation to a metric scale, thereby enabling the use of a much bigger set of analysis techniques and methods.

3.3.4 EGP Classes

The aforementioned scheme for classifying respondents into occupational classes developed by Erikson, Goldthorpe, and Portocarero (1979) is widely known as EGP classes. This classification shows a change in the understanding of class—from skills-based to mutual or reciprocal relations of dependence between employer and employee. When developing the scheme, Erikson and his two colleagues used data from three countries (England, France, and Sweden) to delineate the classes. The data included only men between 20 and 64 years of age. Two main variables were involved: the respondent's current class status at the time of the interview, and the respondent's class origin according to the class status of his father. The classification scheme that eventuated consists of seven different groups, some of which are divided into subgroups, resulting in a total of nine and ranging from proprietors, managers,

administrators, and higher-grade professionals to agricultural workers. As the creators of this class schema admit, problems of distinction are apparent across some of the levels (Erikson et al., 1979).

In the early 1990s, Erikson and Goldthorpe (1992) reworked the scheme, using data from the Comparative Analysis of Social Mobility in Industrial Nations (CASMIN) project. They connected the EGP scheme to the ISCO, a change that requires users to first code the occupation into the ISCO and then adjust it according to information about self-employment and number of employees supervised. The revised scheme consists of 11 classes, ranging from "higher managerial and professional workers" through to "manual supervisors" and "self-employed farmers."

The use of the EGP scheme in education studies poses the challenge of obtaining the information needed. In general, students are highly unlikely to have at hand information on parental self-employment, let alone the number of employees their parents supervise. Therefore, the information needs to be provided by the parents themselves, which necessitates administering a parent questionnaire with the known implication of additional development and administration costs.

3.3.5 Index of Early Home Literacy Activities (EHLA)

As mentioned before, the concept of cultural capital not only concerns possession of cultural items at home, but can exist in three states. One of the three, the embodied state of cultural capital, can be described as the long-lasting dispositions of the mind and the body being inevitably related to the person (Bourdieu, 1986). Bourdieu saw the time investment as the best measure of cultural capital. Following this definition, the researchers involved in some studies attempt to attain information about this embodied state of cultural capital. For example, in the PIRLS 2006 assessment framework (Mullis, Kennedy, Martin, & Sainsbury, 2006), the background context was divided into different aspects of reading literacy. One of these—the home context—consists of a separate aspect called "social and cultural resources." For this aspect, the authors of the framework delineated the necessity of collecting data on the literacy activities of parents with their children that would likely foster positive attitudes among the latter toward reading. These activities, identified as direct guidance, modeling, and support of children's behavior, convey parents' beliefs and attitudes toward literature and reading (Mullis et al., 2006). The framework authors then took the information on the frequency of occurrence of these early literacy activities with the children at home (reading books, telling stories, singing songs, playing with alphabet toys, playing word games, reading aloud signs and labels, etc.) and used it to derive a new index called the Index of Early Home Literacy Activities (EHLA) (Martin et al., 2007, p. 201). These items and the EHLA are meant to measure the building of embodied cultural capital directed toward or related to literature, reading, and language.

Because the information needed relates to the time before children enter school, it is highly likely only the parents will recall if they carried out such activities with their child. Again, additional effort, through the use of a parent questionnaire, is needed

to collect the data. But even then, the longer the span of time that has elapsed since early childhood, the more difficult it is likely to be for parents to recall the activities, resulting in a higher proportion of missing data. In terms of association with learning outcomes or achievement in school, the influence of early home literacy activities will likely wane as the child gets older (i.e., the longer the child attends school), with the most obvious examples of possible influence encompassing schooling.

But there is a special feature of the EHLA. Unlike the other indices mentioned so far, the EHLA is directed toward a specific (and desired) outcome of education—the reading ability of the child. Although reading or reading literacy can be seen as the most basic ability shaping much of a person's living conditions (Mullis et al., 2006), reading-literacy-related activities are clearly goal-oriented and targeted at the child. As such, these activities clearly relate to a specific learning area. Another difference in terms of the family background measures mentioned previously is that the decision of whether or not to engage in these activities is a relatively easy one. In contrast, deciding to change a job usually poses a much higher burden on a person because it can involve uncertainty about future employment and salary, and also because other people need to agree to the change (e.g., to employ the person). Even less self-determined is the increase in one's income, assuming the change in employment eventuates.

3.3.6 The PISA Index of Economic, Social, and Cultural Status (ESCS)

The research team involved in PISA derived an index from various indicators representing economic, social, and cultural status. The composition of the index, originally developed for the first cycle of PISA in 2000, was changed for the 2003 and 2006 cycles (see OECD, 2009, p. 346ff.). The index uses information on home possessions and also takes into account the highest-level occupation and the highest-level education of both parents, expressed as years of schooling. The PISA Index thus consists of the "classical" components of an SES scale (see Chapter 3.3.1). Home possessions has continued to serve in the scale as a proxy for income, because no direct data on income are available.

In the PISA cycles, all the information included in the ESCS index is gathered from student responses. Although a parent questionnaire is offered as an international option, few countries administer it. Because home possessions (as the substitute for the information on income) can be collected from student responses, the index involves few, if any, major challenges for survey administration.

3.4 The Multilevel Nature of Family Background

As already mentioned, family background can be seen as comprising different levels. The usual approach focuses on the individual's family background. Studies that involve the participation of students in school add two levels: the classroom and the school. Indicators at the individual level can also be aggregated to classroom or school level—as averages, for example. This section addresses the potential of taking the different levels into account.

3.4.1 Individual, Classroom, School, and Neighborhood

During the past decades, multilevel techniques for analyzing hierarchical data made great progress. Their need arose due to the fact that individuals are part of a broader system. Students are nested in classes that are nested in schools. This kind of grouping in a hierarchical structure requires analyzing the effect of the higher levels on the lower ones. For example, the "school effect" would refer to the influence of certain school characteristics on student test scores or any other educational outcome at the student level. From the 1960s until the end of the 1980s, the main methods of analyzing school effects encompassed regression methods or analysis of variance. This approach inevitably posed methodological problems such as confounding the variation of the lower and higher levels of the hierarchy. Multilevel methods that were developed later helped overcome these disadvantages. For example, they allowed the variance to be partitioned between the higher and the lower levels, yielded more precise estimates with less bias, and produced more reliable information concerning within and between school effects (Everson & Millsap, 2004).

However, a multilevel or hierarchical structure can be found not only in education systems but also in the social environment of individuals outside school. Each individual is acting in and is a member of multiple environments—families, groups of peers, communities, and neighborhoods. And each one of these environments shapes the individual's behavior, beliefs, and values. For example, the higher the SES of an individual's neighborhood is, the more "extra value" the neighborhood adds to his or her social and material resources (Yang & Gustafsson, 2004, p. 262). Even social institutions treat residents from diverse neighborhoods differently, according them various priorities in varying kinds of social situations. This occurrence also influences education. Children coming from communities with high incomes will most likely attend schools of higher quality and meet other children with similar socioeconomic backgrounds, with the interaction between them likely increasing their occupational and educational aspirations (Yang & Gustafsson, 2004, p. 262).

Usually, children attend schools in their neighborhood. The characteristics of the school will simply reflect the community's characteristics (sociocultural, economic, and ethnic). As Yang and Gustafsson (2004) state, many previous studies suggest the need for multilevel ecological contexts and for investigation of the effects of micro- and macro-levels on educational achievement: "A multilevel structural equation modeling approach can decompose the total variation into different sources (i. e., dimensions) of variations at different levels" (Yang & Gustafsson, 2004, p. 262).

Yang and Gustafsson (2004) identified different dimensions of home-possession items at both the individual and the school level across the countries participating in IEA's Reading Literacy Study (RLS). The authors identified two different dimensions—academic and cultural, and economic and material—at the individual level for the majority of countries, but used different variables for measuring these latent background aspects. Nevertheless, they could identify only a single general capital factor at the within-school level in a good many of the participating countries, probably because the representation of different indicators of cultural capital was

"pitched" too low. At the school level, they identified one single factor representing the community's sociocultural and economic environment. They also found strong correlations between the cultural and economic capital indicators at this level, such that the two aspects of family SES could not be distinguished (Yang & Gustafsson, 2004).

In regard to the importance of the higher levels (classroom, school, and neighborhood context), two major findings need to be noted. First, and in line with the previous statement, data from large-scale education studies show that achievement is also influenced by higher-level factors that are beyond the individual level. For example, Yang (2003), using multilevel methods, found that in about half of the countries he analyzed, approximately 50% of the variation in reading achievement was accounted for by a classroom-level sociodemographic factor. Second, even if the difference in individual-level SES is lower, or even if it does not exist, SES may have a much stronger relationship with achievement at the group level. In their detailed explanation of this phenomenon, Raudenbush and Bryk (2002) argue that the expected difference in the outcomes between two students is actually a contextual effect. The two students might have the same individual SES, yet attend schools with different school SES. Raudenbush and Byrk's analysis (2002, p. 141) shows that students with higher school SES also have higher learning achievement outcomes.

3.4.2 Aggregate Measures

So far, we have discussed mainly cases of individual items serving as background indicators. Sometimes, though, it is not possible to collect particular kinds of information, either because of difficulties students have with answering questions, or because the study's budget does not allow for asking parents, or because asking for the information is seen as too obtrusive. Merola (2005) argues that in such cases data from secondary sources can be obtained, although, of course (and unfortunately), not at the individual level. Merola furthermore argues that because this type of approach was successfully used in the past to measure poverty levels and home prices, it can also be successfully used to obtain other SES data.

Aggregated measures have a number of advantages compared to individual measures. They enable researchers to consider environmental factors that go beyond the individual and influence how the resources are distributed. Furthermore, information can be collected that respondents usually are not willing to provide, such as income, for example. Aggregation can be done at different levels and at levels of different sizes. Aggregate levels can be national level or tracts, block groups, and blocks or postal codes. Given the issue of confidentiality of respondent-supplied information, smaller levels of aggregation are not widely used, although this approach would give more detailed results. Nevertheless, the aggregated-level measures give less precise results compared to the individual ones. Also, smaller levels of aggregation do not reduce the bias significantly (Merola, 2005). Two important aggregated measures of SES, mentioned earlier, are house prices and poverty level.

House Prices

One of the factors closely related with child development is the home the child lives in. Providing a place to live and raise children is not just a matter of the building itself; it also concerns the home's location in the environment that one is willing to reside in. The presence or the absence of various institutions in the vicinity also plays an important role in choosing the family home. Choice, of course, relies on the type of investment families can make. That, in turn, concerns the amount of money that parents are able/willing to pay for a better life, in general, and also (in the context of our current discussion) for educational outcomes. According to Hoxby (2001), house prices are the most prominent kind of this type of investment, at least in the U.S.

Brasington and Haurin (2006) compared school districts of students with a school achievement of one standard deviation below and above the mean. The authors found that house prices for students in the two groups varied by up to about 14%. Higher average levels of achievement were associated with higher house values.

Poverty Level

During the 1990s, Massey and his colleagues (cited in Kurki, Boyle, & Aladjem, 2005) pointed to alternative, aggregated measures of poverty level. Their work led to the Dissimilarity Index, which illustrates the intensity of concentrated poverty, and the Isolation Index, which represents the probability that poor families are in contact only with other poor families. Researchers exploring the influence of poverty on student achievement can draw poverty-level data from census data in two ways: (a) by identifying the poverty level of the neighborhood surrounding the school that the student attends, and (b) by calculating the percentage of single-parent households within the neighborhood. Kurki et al. (2005) found that the poverty level of the school neighborhood is a significant predictor of achievement, although much weaker than other measures such as the free or reduced school lunch program (in the U.S.). Single-parent households have weak predictive power, while the Isolation Index has a strong and negative association with achievement. In general, Kurki et al. (2005) concluded that census-based measures of poverty in a school's neighborhood have lower power, while the Isolation Index, which measures the concentration of poverty, has a strong relationship with achievement.

A general concern regarding these aggregate measures lies with their international comparability. The information pertaining to these measures needs to be collected from additional sources. Information on community level (or any level below the country level) is usually collected at the national or regional level, but most likely only within a country, which raises the question of whether such information is available for all countries participating in an international assessment. Even if there is, the information is not likely to be comparable across countries, because each country independently sets aggregation level, units, and other features of the aggregate measures.

3.5 Administration Issues

In addition to these content-related issues and thoughts, researchers need to consider additional aspects when trying to obtain information about students' family background. Survey design, questionnaire development, sampling, and respondent level are examples of organizational (as opposed to content-related) issues that might affect the quality of the measurement of family background. We briefly introduce these aspects in this section.

3.5.1 Reliability of Student Information about Parents

As already mentioned, information about characteristics of students' parents (such as their occupation) are generally obtained from the students. The question that needs to be asked here is whether student information on parental characteristics is reliable. As research shows (Baratz-Snowden, Pollack, & Rock, 1988), reliability depends, to a large extent, on the age of students: the older they are, the more likely they are to provide the same answer as their parents. A large proportion of students in elementary school do not know or simply do not answer questions concerning parental education.

Referencing data from the U.S.'s High School and Beyond (HS&B) survey, Fetters, Stowe, and Owings (1984) analyzed student against parent responses and found that students provide highly reliable and valid data about the educational attainment of their parents. According to Hauser (1994), the information provided by students at ages 14 and 15 is about as reliable as that given by their parents. Evidence in support of this conclusion is also provided by validation studies regarding the agreement between students and their parents conducted in Canada, the Czech Republic, France, and the United Kingdom during the PISA 2000 cycle. These studies found that data on parental occupation provided by 15-year-old children could be used as a valid indicator of parental occupation (Adams & Wu, 2002, p. 220).

There are, however, other issues regarding the extent of agreement between parents and children on family-background characteristics (see, for example, Baratz-Snowden et al., 1988). One of them concerns the question format itself: open-ended questions about parental occupation yield more valid data than multichoice questions. Another concern is that higher-achieving students provide more valid responses than lower-achieving students. Also, children of higher-educated parents give more accurate information. Issues related to ethnicity also influence the validity of home background information, and female students' responses are generally more valid than male students' responses. Another gender issue is that boys provide more valid information about their father's than their mother's education whereas girls provide more valid information about their mother's than their father's education. This outcome may be because the child identifies most with the parent of the same sex.

The PISA 2006 technical report states that data from the study show a relatively high consistency between student and parent data on questions about education, and a somewhat lower consistency on occupation data (OECD, 2009, p. 57). Schulz (2006) argues that the inconsistency of the responses between students and their parents

from the 15 countries that used the parental questionnaire in the PISA 2006 field trial is due to several reasons: a lack of precision in the instructions for filling out the questionnaire, the tendency to give socially desirable responses, and problems in coding the open-ended responses to questions about education and occupation. Schulz furthermore argues that parents' responses could also be biased, and that this possibility needs to be acknowledged: "Apart from coding problems, lack of precision in the job descriptions and also tendencies to give socially desirable responses might affect the reliability or validity of parent responses" (Schulz, 2006, p. 15). Despite the considerable variation in the reliability of some of the PISA data, Schulz also observed in an earlier work that, in general, student reports about parental occupation are more reliable than the ones about parental education (Schulz, 2005).

3.5.2 Study and Questionnaire Design

International large-scale assessments in education collect huge amounts of data. Family background is only one aspect of interest related to achievement. This fact makes the task of collecting relevant data in a limited time with instruments of a limited length and a limited number of questions a challenge. The difficulty of the task becomes even more apparent when one realizes the breadth and number of the different aspects that such research projects need to cover. Collecting data therefore needs to focus on the most important characteristics, and the survey instrument needs to contain a limited number of items that are measured very precisely. As Hauser (1994, p. 1541) said, the burden of data to be collected should be kept in bounds, and at the same time the focus should be "on characteristics that will be relatively easy to measure, that can be measured for every child in the survey, and that will probably not vary greatly over the short term." The problem associated with the variation in background variables over time and their impact on achievement is influenced by countries' general characteristics.

We need to acknowledge that a universal recommendation about which and how many variables should be included in a questionnaire is not possible. Instead, the selection of variables should be done carefully and with regard to the aim(s) of the study. As Hauser (1994, p. 1545) further noted, "a standard set of racial/ethnic and socioeconomic variables, however well measured, cannot serve as all-purpose statistical controls for family background."

Large-scale assessments in education, as well as in any other area, are very time consuming and expensive. Some of them, such as PIRLS, TIMSS, and PISA, are international comparative studies that collect data from a large number of countries all over the world. Paper and pencil questionnaires serve very well in these huge endeavors of data collection that are conducted within a limited period of time. But at the same time, they have, as is the case with any other measurement tool, some disadvantages. Paper questionnaires handed out to parents in order to obtain information about family background characteristics are self-assessment tools. Self-assessment is subjective. At times, the quality of the instruments can also contribute to the bias. As already mentioned, the PISA 2000 student questionnaires have been criticized with respect to the precision of the occupational job descriptions, which

posed a problem in terms of consistency of the data (Schulz, 2006). Due to restrictions in time and budget, personal interviews with parents are rarely conducted in large-scale assessments in education. Nevertheless, personal interviews could contribute to the quality of the background information collected. But they would also pose a big challenge for survey administration procedures.

Because of our focus in this paper on the TIMSS, PIRLS, and PISA assessments, we now offer a brief overview of the most recent developments in these assessments with regard to the measurement of family background. As trend studies, both TIMSS and PIRLS need to employ, to a large extent, measures and items already used in previous cycles. Not to do so limits ability to report on trends. Nevertheless, there is some room for change or development. The TIMSS 2011 countries that administered TIMSS to Grade 4 students and that administered PIRLS to the same students were able to use the data from the PIRLS home questionnaire (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009). Building on experience gained in previous PIRLS cycles, the PIRLS research team administered the home questionnaire to students' parents. Parents were asked to provide information, such as their occupation, additional to the information on family background gathered from students through the student questionnaire. This approach has enabled TIMSS to report, to some extent, on associations between occupational data and achievement, in line with one of the areas of interest delineated in the TIMSS 2011 framework. The framework's authors set out evidence from the research literature on the strong association between student achievement and parental occupation (Mullis, Martin, Ruddock et al., 2009, p. 114).

PISA relies also on gaining information from students about family background. With the exception of the 2003 cycle, countries participating in PISA can choose to administer a parent questionnaire to the parents of the participating students. Since 2006, PISA has asked parents to provide information on household income and expenditures for educational services (OECD, 2009, p. 60).

IEA's International Civic and Citizenship Education Study (ICCS), which collected data from mainly Grade 8 students in 2009, asked these young people about their knowledge of and attitudes toward democracy, civics, and citizenship. For the first time in recent IEA studies, the research team collected information about the occupation of the students' parents via an open-ended question format that allowed students to enter the job title and a description of that job for both mother and father (or caregivers). Coding of the open-ended answers into an occupation classification scheme was followed by the assignment of ISEI scale scores (see Chapter 3.3.3). In addition, the research team created an Index of Family Socioeconomic Background, using information relating to parental occupation and level of education as well as the number of books at home (see Schulz et al., 2010, p. 222). All scales and indices based on occupation data showed strong relations with civic knowledge, indicating that the effort required to collect this kind of information is worthwhile.

4. Research Questions

As the literature review of family background and approaches to measuring its different aspects has shown, measurement of family background raises a number of issues:

- *The many components to students' family background:* Grasping all of these within one study is a considerable challenge, if even possible, due to different limitations, such as time and financial constraints and limits on the length of the background data-collection instruments.
- *The difficulty of measuring some of the aspects:* This difficulty arises for various reasons, including respondents' unwillingness to provide information on sensitive data, such as income. Other sources of difficulties in obtaining information lie in the lack of census information on important variables such as community SES, the lack of accuracy in answers provided by young students, and answers that respondents make in accordance with what they think are socially desirable answers.
- *Quality:* Even when it is possible to obtain the information, issues remain with regard to quality, in terms of missing data and reliability of the indicators used.
- *Difficulty of operationalizing constructs:* Some of the constructs of family background are quite nebulous and hard to operationalize. Because each study defines different measures of aspects of family background and develops items to measure them in a different way, problems with validity may occur.
- *Different constructs and indicators:* The constructs and indicators selected for inclusion tend to differ across the different studies.
- *Lack of applicability:* Not all of the measures are likely to be applicable in all countries participating in a study, which leads to problems with crosscultural validity.
- *Derivation of indices and scales of latent variables:* The various studies differ in how they derive specific indices or scales of latent variables as part of family background, such as cultural possessions, for example. The studies also use different components of these indices. The composition of the indices can furthermore depend on the research goals of the study.

These issues associated with measuring family background and its different aspects led us to develop and seek answers to the following research questions:

1. Which measures of family background provide the highest quality data in terms of missing data and reliability?
2. How do the quality of scales (in terms of bias, reliability, and validity) and the association with achievement change if components with low quality (in terms of nonresponse and reliability) are removed?
3. Which of the measures of family background and derived variables used in the studies in scope have substantial association with achievement across all participating countries and which of them explain the highest amount of variance in achievement?
4. What are the differences in association between the family background measures and achievement across the studies when controlling for student age groups?
5. Within the context of future large-scale education studies, which family background measures seem to be the most appropriate ones to use when endeavoring to account for students' achievement with regard to different content domains, target populations, missing data, reliability, and validity?

5. Method

The research questions considered in this study relate to different issues associated with family background measures and indices. We addressed them by analyzing the data from three large-scale assessment studies, using the methods detailed in this chapter of our paper. We describe the data from the different studies that we analyzed as well as the statistical methods we used to analyze the various issues related to measures of family background.

5.1 Data

All datasets used in the analyses were taken from international large-scale assessments conducted in countries from all over the world. The studies, all well known, are the following:

- The IEA Trends in International Mathematics and Science Study (TIMSS);
- The IEA Progress in International Reading Literacy Study (PIRLS); and
- The OECD Programme for International Student Assessment (PISA).

All three studies are repeatedly conducted in cycles of three (PISA), four (TIMSS), or five years (PIRLS). We used data from the most recent cycles—that is, TIMSS 2007, PIRLS 2006, and PISA 2006—because they reflect the latest methodological developments. Most of the data considered is drawn from student responses. Where available, we also analyzed data from the students' parents.

The following sections provide short descriptions of these studies as well as of the variables within the instruments (the student questionnaire and parent/home questionnaire) used in the studies that relate to family background. Also considered are derived variables or scales of family background composed from single measures reported in the studies' findings. We end this section with a description of the set of countries we included in our analyses.

5.1.1 Studies

As already stated, the current research employed data from TIMSS 2007, PIRLS 2006, and PISA 2006. The description of each study that follows includes introductions to the variables and derived indices of family background used in them.

TIMSS 2007

The Trends in International Mathematics and Science Study (TIMSS) is a large-scale assessment survey designed to compare students' achievement in mathematics and science (as the name itself shows) in many countries around the world. TIMSS is a trend study conducted by the International Association for the Evaluation of Educational Achievement (IEA) and is directed by the TIMSS and PIRLS Study Center located in the U.S. at Boston College. The study is conducted in a four-year cycle, with the initial study taking place in 1995 and subsequent cycles in 1999, 2003, and 2007. The study targets two student populations—students in Grade 4 and students in Grade 8.

The total number of countries participating in TIMSS 2007 was 59, with an additional eight benchmarking participants.[2] The number of participating countries at fourth grade was 37 countries plus seven benchmarking participants, with over 183,000 participating students in total. Fifty countries and seven benchmarking participants took part in the survey of eighth-grade students, who totaled more than 241,000 in number (Martin et al., 2008).

In addition to collecting achievement data in the two content domains (mathematics and science), TIMSS 2007 also collected data on the social and educational contexts of students related to curriculum, schools, teachers and their preparation, and classroom activities and characteristics. These contextual data were collected by means of four different questionnaires:

1. *The curriculum questionnaire*, completed by the TIMSS national research coordinators (NRCs) of the participating countries;
2. *The student questionnaire*, completed by the students;
3. *The teacher questionnaire*, completed by the students' teachers and administered in three forms—one for the fourth-grade class teacher and two different ones for the science and mathematics teachers of the eighth–grade students; and
4. *The school questionnaire*, completed by the principals of the schools of the students surveyed (Mullis et al., 2005).

Because our focus in this paper is on the family backgrounds of students, we used only the student questionnaire data from both target populations (Grades 4 and 8). Each student participating in the mathematics and science assessment completed the student questionnaire. It contained questions concerning students' circumstances at home and school, classroom experiences, self-perceptions of aptitude for and attitudes toward both subjects of the assessment, homework, out of school activities, computer use, and basic demographic information (Mullis et al., 2005). TIMSS does not collect data from parents.

The TIMSS 2007 student questionnaire data on family background included the following aspects: the frequency with which the language of the TIMSS assessment

2 Benchmarking participants are "regional entities that follow the same assessment procedures as the countries;" see Olson, Martin, and Mullis (2008, p. 2).

was spoken at home, the number of books at home, the availability of specific educational resources at home, the availability of specific home possessions, the immigration status of the student and each one of his or her parents, and the highest level of education of each of the parents. Note, however, that the information on parental education was collected from the Grade 8 students only. For more details on the items and variables pertaining to family background in TIMSS 2007, see Table 5.1.

Table 5.1: Items and variables in the TIMSS 2007 student questionnaire associated with family background

Item content	Description	Variable Grade 4	Variable Grade 8
Language	Student's frequency of use of the language of the test at home	AS4GOLAN	BS4GOLAN
Books in the home	Number of books in the student's home	AS4GBOOK	BS4GBOOK
Home possessions	Several items about educational resources and general possessions in the student's home*		
	Calculator	AS4GTH01	BS4GTH01
	Computer	AS4GTH02	BS4GTH02
	Study desk	AS4GTH03	BS4GTH03
	Dictionary	AS4GTH04	BS4GTH04
	Internet connection	AS4GTH05	BS4GTH05
Parents born in country	Whether or not mother was born in country	AS4GMBRN	BS4GMBRN
	Whether or not father was born in country	AS4GFBRN	BS4GFBRN
	Whether or not parents were born in country	ASDGBORN	BSDGBORN
Student born in country	Whether or not student was born in country	AS4GBORN	BS4GBORN
	and, if not, age at which student emigrated	(AS4GBRNC)	(BS4GBRNC)
Parents' education (Grade 8 only)	Highest level of education completed by father		BS4GMFED
	Highest level of education completed by mother		BS4GFMED
	Highest level of education completed by parents		BSDGEDUP

Note: * In addition to the five home-possession items included in the international questionnaire, TIMSS offered participating countries the opportunity to survey country-specific items. Because these differed from country to country, they are not included in our analyses.

In addition to containing information on single-indicator variables, the TIMSS 2007 international database also contains derived variables using single measures from the student questionnaire. Because TIMSS surveys two different populations (Grade 4 and Grade 8 students), there are separate derived family background variables for each. Only one family background variable is derived from the Grade 4 student questionnaire—"parents born in country." It is based on the students' responses to two separate questions that ask students whether their mothers and fathers were born in the country where they (the students) were taking the assessment. The source variables are dichotomous (yes/no), and the values in the derived variable depend on combinations of students' answers about the country of birth of each parent. The derived variable has three categories: both parents born in country, only one parent born in country (students respond with "yes" for one of the parents and "no" or missing answer for the other), and neither parent born in country ("no" for both parents or missing answers for both parents) (Foy & Olson, 2009a, p. 6).

There were two derived variables for the second TIMSS 2007 population (Grade 8 students). One of them was again "parents born in country," which was computed in the same way as described above for the first population (Grade 4 students). The other variable derived from the Grade 8 student questionnaire was "parents' highest education level," taken from students' responses to the question about the highest level of education attained by each parent. If data for the mother's highest level of education were missing, the value from the father's education was used, and vice versa. If data about both parents were missing, the derived variable was set to missing. If both variables had valid data, the value for the parent with the higher educational level was taken (Foy & Olson, 2009a, p. 56).

PIRLS 2006

The Progress in International Reading Literacy Study (PIRLS) is also an international large-scale trend assessment conducted by IEA and directed by the International Study Center at Boston College. PIRLS assesses reading literacy in cycles of five years. The first PIRLS cycle took place in 2001, and the second one was conducted in 2006. PIRLS assesses the reading literacy of students in Grade 4.

PIRLS 2006 was conducted in 40 countries around the world. Belgium participated with two education systems, and five Canadian provinces took part independently, making for a total of 45 participating education systems (Mullis et al., 2007, p. 1). The total number of tested students was over 200,000 (Foy & Kennedy, 2008a, p. 1).

PIRLS 2006 also used contextual questionnaires, with the aim of gathering information on four different contexts to aid understanding of the achievement data. These contexts were national and community, home, school, and classroom. The vehicles used to collect this information were the following:

1. The student questionnaire, completed by the participating students;
2. The so-called Learning to Read Survey (also known as the home questionnaire), completed by the parents of each participating student;
3. The teacher questionnaire, completed by the students' respective teachers;
4. The school questionnaire, completed by the school principals; and
5. The curriculum questionnaire, completed by each participating country's NRC (Mullis et al., 2006).

Each student taking part in the 2006 study completed the student questionnaire. It collected information on various student characteristics: the student's home and school life (classroom experiences, reading as homework), the student's perception of his or her ability as a reader and his or her attitudes toward reading, the frequency with which the student read out of school and used computers, and the type of literacy resources at home. The student questionnaire also included items designed to collect basic demographic information (Mullis et al., 2006).

In contrast to TIMSS 2007, PIRLS 2006 also used a parent questionnaire (the aforementioned Learning to Read Survey). The parents of each student taking part in the reading achievement test were asked to complete it. The questionnaire included

items for gathering information from parents (or primary caregivers) about literacy interactions between parents and child, literacy resources in the home, parents' reading habits and attitudes toward reading, and the connections between the parents and their child's school. The questionnaire also collected basic demographic and socioeconomic information (Mullis et al., 2006).

The PIRLS 2006 student questionnaire and Learning to Read Survey collected data on different aspects of family background. The items in the student questionnaire that sought out this information were those on the frequency with which the language of the test was used in students' homes, the number of books at home, the availability of particular educational resources and home possessions, and the immigration status of each student and of each of his or her parents.

The Learning to Read Survey (or "home" questionnaire) contained items on the language of children's books in the home, the language most often used at home, the number of books at home, the level of education of each parent, and the employment situation and main job of each parent. The questionnaire also contained a measure of the parents' financial wellbeing (as perceived by the parents). For more detailed information on the questions included in the student and home questionnaires and their associated variables, refer to Tables 5.2 and 5.3.

Table 5.2: Items and variables in the PIRLS 2006 student questionnaire associated with family background

Item content	Description	Variable
Language	Student's frequency of use of the language of the test at home	ASBGLNGH
Books in the home	Number of books in the student's home	ASBGBOOK
Home possessions	Several items about educational resources and general possessions in the student's home*	
	Computer	ASBGTA1
	Study desk	ASBGTA2
	Own books	ASBGTA3
	Daily newspaper	ASBGTA4
	Own room	ASBGTA5
	Own cellphone	ASBGTA6
Parents born in country	Whether or not mother was born in country	ASBGBRNM
	Whether or not father was born in country	ASBGBRNF
	Whether or not parents were born in country	ASDGBRN
Student born in country	Whether or not student was born in country	ASBGBRN1

Note: * In addition to the six home-possession items included in the international questionnaire, PIRLS offered participating countries the opportunity to survey country-specific indicators of wealth. Because these differed from country to country, they are not included in our analyses

Table 5.3: Items and variables in the PIRLS 2006 Learning to Read Survey associated with family background

Item content	Description	Variable
Language of books	Language of children's books in the home	ASBHCHBL
Language spoken	Language used most often when parents speak with their child (F=father, M=mother)	ASBHLAHF ASBHLAHM
Books in the home	Number of books in the home	ASBHBOOK
Children's books in the home	Number of children's books in the home	ASBHCHBK
Parents' education	Highest level of education completed by father Highest level of education completed by mother Highest level of education completed by parents	ASBHLEDF ASBHLEDM ASDHEDUP
Parents' employment situation	Employment situation of father Employment situation of mother Employment situation of parents	ASBHEMPF ASBHEMPM ASDHPEMP
Parents' main job	Father's main job Mother's main job Parents' occupational level	ASBHMJF ASBHMJM ASDHOCCP
Family financially well off	Self-report of family's financial situation relative to other families	ASBHWELL
Early home-literacy activities	Reading books Telling stories Singing songs Playing with ABC-tools Playing word games Reading aloud signs and labels	ASBHHA01 ASBHHA02 ASBHHA03 ASBHHA04 ASBHHA07 ASBHHA09

In addition to the particular variables measuring different family characteristics, PIRLS derived variables that were computed by combining data from single items. Derived variables "provided a more comprehensive picture of the construct of interest than the individual variables could on their own" (Martin et al., 2007, p. 197). Not all students were included in the computation of the derived variables: those who had missing values for a certain number of the single items included in the calculation were assigned a missing value for the derived variables. The indices (derived variables) in the PIRLS 2006 international database consisted of three categories—high, medium, and low. When constructing an index of the components (single items), researchers need to intercorrelate these in order to produce a reliable scale (Martin et al., 2007).

PIRLS 2006 has two family background indices of interest within the context of our study—the Index of Home Educational Resources and the Index of Early Home Literacy Activities—as well as four other derived variables: "student's parents born in country," "parents' highest education level," "parents' employment situation," and "parents' highest occupational level."

The Index of Home Educational Resources (HER) is based on the responses to the questions in both the student and the parent questionnaire about the educational resources at home. The index "is intended to summarize the students' and parents' reports about aspects of the home environment and the extent to which it supports

literacy" (Martin et al., 2007, p. 201). When deriving this variable, the research term used the following variables from both instruments (Foy & Kennedy, 2008c; Martin et al., 2007):

- Number of books in the home;
- Number of children's books in the home;
- Availability of home-possession items (four variables for different items at home—computer, study desk, own books, daily newspaper); and
- Highest level of educational attainment of either parent.

Interested readers might like to look at the original questions from the questionnaires published in Supplement 1 of the PIRLS 2006 user guide (Foy & Kennedy, 2008c). The user guide also includes the response categories of each one of the variables as well as the categories of the Index of Home Educational Resources itself, which is Supplement 3 of the guide.

The second index of family background that was of interest to us was the Index of Early Home Literacy Activities (EHLA). This index was obtained from parents' responses about the frequency of engaging in different reading-related activities with their children before their children started primary school. The index places students into three categories—high (2.33 through 3), medium (1.67 through 2.33) and low (1 to less than 1.67), with the numbers referring to the average that was computed across the six items based on a three-point scale: "never or almost never" = 1, "sometimes" = 2, and "often" = 3 (Martin et al., 2007).

The separate variables used to construct this index were the six items of the question, "Before your child began primary school, how often did you or someone else in your home do the following activities with him or her?" Each item had three response categories—"often", "sometimes," and "never or almost never." The six items included reading books, telling stories, singing songs, playing with alphabet toys, playing word games, and reading aloud signs and labels (Foy & Kennedy, 2008c). For more details on the Index of Early Home Literacy Activities, see the home questionnaire in Supplement 1 and the full description of the index in Supplement 3 of the PIRLS 2006 user guide for the PIRLS 2006 international database (Foy & Kennedy, 2008c).

Another derived variable is "student's parents born in country," obtained from questions in the student questionnaire that asked whether the mother and the father were born in the country of the literacy achievement test. The derived variable has three response categories: "both" (if the student responded that both his or her mother and father were born in the country of the test), "either" (one parent born in the country of the test and the other in a different country), and "neither" (both parents born in a country other than the one of the test) (Foy & Kennedy, 2008c, p. 4).

The variable "parents' highest education level" was derived from questions in the home questionnaire concerning the highest level of education completed by both parents. The variables used for the computation were first collapsed by merging the categories concerning postsecondary nontertiary education into one category and

merging the categories for higher education into another. The value for the parent with the highest education level was taken and all "not applicable responses" were recoded as missing (Foy & Kennedy, 2008c, p. 22).

"Parents' employment situation" was derived using the home questionnaire item that asked about the employment status of both parents. The derived variable was based on the answers of the parents and had the following categories: both fulltime; either fulltime; both less than fulltime (for parents answering that they worked part-time for money or that they did not work in any type of paid job, but were looking for such a job); and other (for those who responded that they had some other kind of working arrangement or answered that none of the response categories was applicable to them) (Foy & Kennedy, 2008c, p. 23).

"Parents' highest occupational level" was derived from the answers of the parents to questions about the kind of work each one of them was doing. The response categories of the source variables about the professions were collapsed from 12 to 7 in the derived variable (from "professional" to "never worked outside of home for pay" plus "not applicable") (Foy & Kennedy, 2008c, p. 24).

PISA 2006

The 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" (OECD, 2009, p. 3). PISA was launched in 1997 by the Organisation for Economic Co-operation and Development (OECD). This was done, as the founders of PISA stated, "In response to the need for cross-nationally comparable evidence on student performance" (OECD, 2007, p. 3).

PISA surveys the key competencies of 15-year-old students in order to answer the question of whether these students are prepared to participate in the global economy and real-life challenges after finishing their schooling. PISA does not focus on a specific grade level in school. Rather, it defines its target population as students between 15 years and 3 months of age and 16 years and 2 months of age at the time of the assessment. In addition, students who take the test should have completed at least six years of formal schooling. This age bracket was selected to capture students (in the participating countries) just as they are about to finish compulsory education (OECD, 2007, pp. 22–23).

PISA is conducted in cycles of three years, the first of which took place in 2000, and the following in 2003 and 2006. Each cycle has a major emphasis on one of the subject domains PISA assesses (mathematics, science, and reading). The PISA assessment of 2006 focused on science (OECD, 2009, p. 28). The total number of countries that participated in this assessment was 57, 30 of which were OECD countries and 27 of which were OECD partner countries (OECD, 2007, p. 17). Nearly 400,000 students took part in the survey (OECD, 2009, p. 22).

In addition to administering the achievement tests, PISA 2006 used contextual questionnaires—a student questionnaire, a school questionnaire, an information and communication technologies (ICT) familiarity questionnaire (completed by students), and a parent questionnaire. The last two questionnaires were optional, that is, countries were free to decide whether to use them or not (OECD, 2009, p. 59). The parent questionnaire was administered in 10 OECD and six partner countries only (OECD, 2009, p. 24).

Besides collecting information on basic student characteristics, the PISA 2006 student questionnaire asked about different family background aspects: parental occupation and education, home possessions and number of books at home, immigration status of both parents and the student, and the language most commonly used at home (OECD, 2009, p. 58). The parent questionnaire included items on parental background: age, occupation, and education of each parent, and total income of the household (OECD, 2009, p. 60). For more details on the student and parent questionnaire items, refer to Tables 5.4 and 5.5.

Along with the single items in these two background instruments, PISA 2006 employed derived variables or scales, obtained by combining the information from the questions in the student and parent questionnaires. The family background indices computed from individual items taken from the student questionnaire included the following:

- Home possessions;
- Cultural possessions;
- Home educational resources;
- Family-wealth possessions;
- Occupational status of parents;
- Indices of blue-collar/white-collar parental occupation;
- Indicators of a science-related career of both parents and science-related career expectations of students;
- Indices of parental education;
- Immigration status of the mother, father, and student; and
- Language most commonly spoken at home.

One additional index was calculated on the basis of three other derived variables. This was the Index of Economic, Social, and Cultural Status (ESCS), which was computed from three derived variables—home possessions, parents' estimated years of schooling, and highest occupational status of parents. The home possessions, cultural possessions, home educational resources, and family-wealth possessions scales are all indices derived from home possessions (OECD, 2009, pp. 304ff.).

Table 5.4: Items and variables in PISA 2006 student questionnaire associated with family background

Item content	Description	Variable	Scale
Language	Language spoken at home	ST12Q01	
Books in the home	Number of books in the student's home	ST15Q01	
Home possessions	Several items about educational resources and general possessions in the student's home*		WEALTH, CULTPOS, HEDRES, HOMEPOS (using different combinations of the home-possession variables)
	Study desk	ST13Q01	
	Own room	ST13Q02	
	Quiet place to study	ST13Q03	
	Computer for school work	ST13Q04	
	Educational software	ST13Q05	
	Link to internet	ST13Q06	
	Own calculator	ST13Q07	
	Classic literature	ST13Q08	
	Books of poetry	ST13Q09	
	Works of art	ST13Q10	
	Books to help with school work	ST13Q11	
	Dictionary	ST13Q12	
	Dishwasher	ST13Q13	
	DVD/VCR player	ST13Q14	
	Cellphone (number of)	ST14Q01	
	Television (number of)	ST14Q02	
	Cars (number of)	ST14Q03	
	Rooms with bath or shower (number of)	ST14Q04	
Parents born in country	Whether or not mother was born in country	ST11Q02	IMMIG
	Whether or not father was born in country	ST11Q03	
	Mother's country of birth	COBN_M	
	Father's country of birth	COBN_F	
Student born in country	Whether or not student was born in country	ST11Q01	
	(Student's age at time of immigration)	(ST11Q04)	
	Student's country of birth	COBN_S	
Occupation of parents	Occupation of mother ISEI coded	BMMJ	
	Occupation of father ISEI coded	BFMJ	
	Highest occupation of parents ISEI coded	HISEI	
	Mother white-collar/blue-collar classification	MSECATEG	
	Father white-collar/blue-collar classification	FSECATEG	
	Highest parent white-collar/blue-collar classification	HSECATEG	
Education of parents	Highest education of mother	MISCED	PARED (years of schooling)
	Highest education of father	FISCED	
	Highest education of parents	HISCED	

Note: *In addition to the home possession items included in the international questionnaire, PISA offered participating countries the opportunity to survey country-specific items. Because these differed from country to country, we did not include them in our analyses.

Table 5.5: Items and variables in PISA 2006 parent questionnaire associated with family background

Item content	Description	Variable
Household income	Relative to median	PA15Q01
Occupation of parents	Occupation of mother ISEI coded	PQBMMJ
	Occupation of father ISEI coded	PQBFMJ
	Highest occupation of parents ISEI coded	PQHISEI
Education of parents	Highest education of mother	PQMISCED
	Highest education of father	PQFISCED
	Highest education of parents	PQHISCED

Two more groups of indices were computed from variables taken from the parent questionnaire: the Educational Level of Parents, and the Index of Occupational Status of Parents (OECD, 2009, pp. 309–310). Both indices include three separately derived variables—one for the mother, one for the father, and one for the highest educational level/occupational status of both parents.

The different scales on home possessions in PISA 2006 use different sets of items about possessions available at home. The Home Possessions Scale is a summary index derived from items on home possessions in the student questionnaire. These related to the availability of one's own room, a study desk, internet connection, and so on. Another set of questions asked about the number of different items available at home (cellphones, televisions, computers, cars, and rooms with bath or shower). One more variable—number of books at home—was added to the calculation of the Home Possessions Scale (OECD, 2009, p. 316).

The scale was constructed in two steps. First, item parameters for the aforementioned items were estimated for each country based on the item set. The sum of parameters was then constrained to zero for each country. Second, the item parameters were anchored, and the country-specific items were then appended. Each country was scaled separately (OECD, 2009, p. 317).

The cultural possessions and home educational resources scales were derived in a similar way, but only in one step, thereby allowing the item parameters to vary by country. The cultural possessions scale was calculated using three items from the student questionnaire—availability of classic literature, books at home, and works of art at home. The home educational resources scale uses seven items from the home-possessions questions included in the student questionnaire—study desk, place to study, computer to use for school work, educational software, own calculator, books helpful for school work, and dictionary (OECD, 2009, pp. 316–317).

The family wealth possessions scale uses items on home possessions included in two questions from the student questionnaire. The first question is about whether or not the student had a room of his or her own at home and whether the following items were also available in the home: an internet connection, a dishwasher, a DVD player, and three country-specific items. The second question asked about the number

of cellphones, televisions, computers, and cars in the home. The family wealth possessions scale was constructed in the same way as the Home Possessions Scale (OECD, 2009, pp. 316–317).

The occupational data for both parents was obtained through open-ended questions in the student questionnaire that were then coded into ISCO-88 codes. These codes, in turn, were mapped to the International Socio-Economic Index of Occupational Status (ISEI). Three such variables with ISEI scores were produced: one for the mother, one for the father, and one indicating the highest ISEI score of either parent (OECD, 2009, p. 305).

The Index of Blue-Collar/White-Collar Parental Occupation was based on students' reports on their parents' jobs. The ISCO-88 codes were recoded into four categories (blue collar, white collar, high skilled, low skilled) and set up as two separate variables for the mother and father employment categories. An additional variable was then created from these two variables, namely the highest employment category of either parent (OECD, 2009, p. 306).

Another set of variables derived from the information on occupation concerned the science-related occupations of parents and the type of occupation students hoped to pursue in the future. These variables were computed by aggregating ISCO-88 codes into two categories, one called science-related occupations, and the other called no science-related occupations/undetermined (OECD, 2009, p. 306).

Information on parental education from the student questionnaire was recoded (from 0 = none through to 6 = ISCED 5A & 6) into variables indicating the educational level of the mother and the father. In addition, the Index of Highest Educational Level was obtained by taking the highest ISCED level of either parent. The completed education level was converted into the estimated number of years of schooling for both parents (OECD, 2009, p. 305).

The original questions asking whether the mother's, father's, and student's country of birth was the same as the country in which the student completed the PISA assessment were copied into new variables and then recoded into dichotomous variables (i.e., country of birth is the same as the country of assessment or the opposite). The Index of Immigrant Background, calculated on the basis of the original values of the copied variables indicating country of birth, had three categories: native students (students with at least one parent born in the country), first-generation immigrant student (born outside the country with parents also born outside), and second-generation student (born in country, but parents born outside). Students who did not respond to these questions, either about themselves or both parents, were assigned missing values for this variable (OECD, 2009, pp. 30–306).

The language spoken at home variable was derived from a question in the student questionnaire about languages spoken at home and was then recoded into three categories: same as the language of assessment; language of the country, but different from the one of the assessment; and foreign language (OECD, 2009, p. 306).

The Index of Economic, Social, and Cultural Status (ESCS) was derived from three other indices—highest occupational status of parents, estimated years of parental schooling, and home possessions. "The ESCS scores were obtained as component scores for the first principal component with zero being the score of an average OECD student and one the standard deviation across equally weighted OECD countries" (OECD, 2009, p. 346). The ESCS for the OECD partner-countries was derived by summing the products of the OECD factor loadings and the OECD-standardized highest occupational status of parents, estimated years of schooling of parents, and home possessions, and then dividing the result by the eigenvalue of the first principal component (OECD, 2009, p. 346).

The Educational Level of Parents Index was the first of the group of indices derived from the parent questionnaire in PISA 2006. The questions for the educational level of the mother and of the father in the parent questionnaire contained four options; parents were asked to check (tick) the ones that applied to them. The educational level for each parent was obtained by taking only the answer from the parent who had the highest completed level of education. The Highest Educational Level of Parents Index therefore represents the highest completed level of one parent, not both (OECD, 2009, p. 309).

The Occupational Status of Parents Index was derived from the occupation information in the parent questionnaire, and it was obtained from the information included in the open-ended questions. The parents' responses were coded into ISCO-88 codes, recoded into ISEI values, and the scores for mother's occupation and father's occupation were then produced (two variables—one for each parent). The Highest Occupational Level of Parents, therefore, also represents the answer from the parent with the highest occupational level (OECD, 2009, pp. 309–310).

5.1.2 Countries

Part of our research project has been concerned with evaluating measures of family background with regard to different purposes, a circumstance that has implications for administering the studies of interest in different sets of countries. Different studies collect background information on different aspects and in a different way. It is therefore necessary to verify whether and to what extent certain variables or scales derived from those variables are associated with achievement. We could expect that two different studies measuring the same background characteristics would compute nearly the same association between the background measures and achievement. If not, then the reasons for the differences should be sought. However, such comparisons have two problems. First, the large-scale studies described so far are conducted in different countries. Second, the PIRLS 2006 assessment had a target population of Grade 4 students, PISA 2006 surveyed students between 15 and 16 years of age, and TIMSS 2007 collected data on students from Grades 4 and 8. It is fair to assume that students in such different age groups will have different knowledge about their families' background characteristics, not to mention ability to produce accurate information. For example, and as we discussed earlier in this paper, studies show that students' knowledge about their parents can differ considerably

across different age groups. In general, the older the students are, the more accurate their answers about background characteristics are likely to be (Baratz-Snowden et al., 1988; Ensminger et al., 2000; Hauser, 1994).

To address the problem of accuracy in responses from students in different age groups, we conducted two comparisons—one between the PIRLS 2006 and TIMSS 2007 (Grade 4) results, and the other between the PISA 2006 and TIMSS 2007 (Grade 8) results. This separation allowed us to control for student age (with regard to accuracy of their knowledge) when endeavoring to determine the association between home background measures and achievement.

To address the problems associated with different countries participating in the studies, we compared the results for those countries that participated in the pairs of studies that targeted students in the same or near-same age groups, thus PIRLS 2006 and TIMSS 2007 Grade 4, and PISA 2006 and TIMSS 2007 Grade 8. Some of the countries that participated in TIMSS 2007 conducted the assessment with both student populations; others in only one. In order to match the countries with data from both studies in each pair, we used the information contained in Table 5.6. The countries that participated in the two pairings can be easily identified from the table.

Table 5.6 shows the countries participating in the studies in scope. Some of the countries that participated in TIMSS 2007 conducted the assessment with both student populations; others in only one.

5.2 Analysis

To assess the quality of the indicators of family background, we analyzed the data with regard to the following criteria: nonresponse, association with achievement, and, in the case of scales, reliability. Using as our basis the empirical findings from the analysis, we categorized the criteria to aid ease of comparison. We used SPSS to compute response rates and reliability statistics, and the IEA IDB Analyzer software to calculate the association of family background indicators with achievement. The latter included the calculation of standard errors (SEs). In order to retain readability of the result tables, SEs were not included in the tables.

5.2.1 Nonresponse

Every study usually has cases in which some respondents do not provide answers to certain questions. Missing data originating from nonresponse can be a serious problem in research. One of the issues with nonresponse is that it decreases the size of the effective sample drawn from the population of interest.

Reasons for nonresponse vary: refusal to answer, accidental skipping of items, lack of interest in the issue or topic, and lack of knowledge needed to answer the question. In research, nonresponse causes problems because it is expected that each record in the dataset has values on all variables. The usual and easiest solution to this problem is to exclude cases that have missing values for any of the variables included in an analysis (Alison, 2002).

Table 5.6: Countries participating in PIRLS 2006, PISA 2006, and TIMSS 2007

Country/ education system	PIRLS 2006	TIMSS 2007 (Grade 4)	PISA 2006	TIMSS 2007 (Grade 8)	Country/ education system	PIRLS 2006	TIMSS 2007 (Grade 4)	PISA 2006	TIMSS 2007 (Grade 8)
Algeria		•		•	Liechtenstein			•	
Argentina			•		Lithuania	•	•	•	•
Armenia		•		•	Luxembourg	•		•	
Australia		•	•		Macao-China			•	
Austria	•	•	•		Macedonia, Republic of	•			
Azerbaijan			•		Malaysia				•
Bahrain				•	Malta				•
Belgium			•		Mexico			•	
Belgium (Flemish)	•				Moldova, Republic of	•			
Belgium (French)	•				Mongolia		•		•
Bosnia and Herzegovina				•	Montenegro			•	
Botswana				•	Morocco	•	•		•
Brazil			•		Netherlands	•	•	•	
Bulgaria	•		•	•	New Zealand	•	•	•	
Canada			•		Norway	•	•	•	•
Canada, Alberta	•				Oman				•
Canada, British Columbia	•				Palestinian National Authority				•
Canada, Ontario	•				Poland	•		•	
Canada, Québec	•				Portugal			•	
Chile			•		Qatar	•	•	•	•
Chinese Taipei	•	•	•	•	Romania	•		•	•
Colombia		•	•	•	Russian Federation	•	•	•	•
Croatia			•		Saudi Arabia				•
Cyprus				•	Scotland	•	•		•
Czech Republic		•	•	•	Serbia			•	•
Denmark	•	•	•		Singapore	•	•		•
Egypt				•	Slovak Republic	•	•	•	
El Salvador		•		•	Slovenia	•	•	•	•
England	•	•		•	South Africa	•			
Estonia			•		Spain	•		•	
Finland			•		Sweden	•	•	•	•
France	•		•		Switzerland			•	
Georgia	•	•		•	Syrian Arab Republic				•
Germany	•	•	•		Thailand			•	•
Ghana				•	Trinidad and Tobago	•			
Greece			•		Tunisia		•	•	•
Hong Kong SAR	•	•	•	•	Turkey			•	•
Hungary	•	•	•	•	Ukraine		•		•
Iceland	•		•		United Kingdom			•	
Indonesia	•		•	•	United States	•	•	•	•
Iran, Islamic Republic of	•	•		•	Uruguay			•	
Ireland			•		Yemen		•		
Israel	•		•	•	*Benchmarking participants*				
Italy	•	•	•	•	Alberta, Canada		•		
Japan		•	•	•	Basque Country, Spain				•
Jordan			•	•	British Columbia, Canada		•		•
Kazakhstan		•			Dubai, UAE		•		•
Korea, Republic of			•	•	Massachusetts, USA		•		•
Kuwait	•	•		•	Minnesota, USA		•		•
Latvia	•	•	•		Ontario, Canada		•		•
Lebanon				•	Québec, Canada		•		•

This practice, however, introduces the problem of sample size mentioned above. As McKnight and his colleagues (McKnight, McKnight, Sidani, & Figueredo, 2007) argued, the biggest problem associated with nonresponse is its influence on the results of a study. If there is a discernible pattern of missing responses (e.g., one group of respondents with certain characteristics in the sample has mostly missing data on the variables of interest), this will lead to overestimating the information provided from the respondents who do not share these characteristics. In general, the greater the amount of missing data, the larger the impact will be on ability (reduced) to generalize the study findings and to draw statistical inferences.

In addition, decreasing the sample size can result in loss of representativeness, and nonresponse can lead to biased estimates and wrong statistical conclusions (McKnight et al., 2007, pp. 6–7). Moreover, each study tries to achieve high-quality results in terms of reliability and validity. Missing data influence the reliability and validity of the indicators and constructs as well as ability to generalize findings (McKnight et al., 2007, pp. 11–13).

These are all reasons why we wanted to analyze the amount of nonresponse to family background variables in the three studies of interest. Missing data become an especially important issue when indicators of family background are used to derive variables, scales, and indices representing some latent characteristic.

The analyses that follow compare the amount of nonresponse per family background variable, scale, or index composed from separate family background indicators across the three studies of interest. We conducted a separate analysis for each study, and then conducted comparative analyses for each pair of studies (PIRLS 2006 and TIMSS 2007 Grade 4, and PISA 2006 and TIMSS 2007 Grade 8) as stated in the previous section.

Because we are concerned in this paper with comparing measures of family background across the three large-scale assessment studies, we set the criteria that we used to classify the quality of data with regard to nonresponse not solely according to theoretical considerations, but also according to empirical findings. Although a certain amount of nonresponse (e.g., 7% as just an arbitrary example) could be seen as still providing sufficient information for analysis purposes, it might also be seen as a moderate amount when compared to nonresponse rates with other indicators from the studies in our scope. Therefore, for the purpose of our analyses, we categorized a nonresponse rate below 5% as "low," a nonresponse rate between 5 to 15% as "moderate," and a rate above 15% as "high."

In addition to providing the median nonresponse across all of the participating countries except the benchmarking participants, we included the inter-quartile range (IQR) as a measure of dispersion. This measure indicates the variation of the nonresponse rate for the middle 50% of the countries, that is, the difference between the nonresponse rate of the upper end of the lowest 25% of countries and the nonresponse rate of the lower end of the upper 25% of countries. The IQR thus indicates variation across countries. A low IQR represents similar nonresponse across

countries, whereas a high IQR points to considerable differences in nonresponse rates across countries. With regard to defining thresholds for categorizing the IQR for nonresponse as "low," "moderate," or "high," there is no universally valid or agreed upon default value. Rather, the distribution of nonresponse rates needs to be checked if the IQR indicates substantial differences across countries.

According to the design of the three studies of interest, data are collected from parents only if their children participated in the assessment. Thus, unit nonresponse might occur with regard to data from the home/parent questionnaire. Consequently, nonresponse for student questionnaire data is a matter of item nonresponse only (i.e., a student did not answer a particular item or question), whereas nonresponse for home/parent questionnaire data is the combination of item and unit nonresponse (i.e., parents who did not return the home/parent questionnaire at all).

5.2.2 Association with Achievement

One major purpose of TIMSS, PIRLS, and PISA is to measure achievement in specific subject content domains—mathematics, science, and reading. As we have explained, all three studies collect background information about students and their parents, teachers, and school. This background information is important for the analysis of contextual factors (resources and activities) that can foster achievement (Mullis et al., 2006). As Sirin (2005), for example, points out, SES is both directly and indirectly linked to academic achievement. The indirect link is through multiple interacting systems.

Because the influence of different family background components on achievement has been well documented in the research literature, we wanted to determine the strength of the association of background variables (including single components as well as derived variables) with students' overall achievement per content domain by computing correlation coefficients and then analyzing the explained variance using simple linear regression. For PIRLS 2006, we investigated the association of the background variables with achievement, that is, the overall reading achievement score. With TIMSS 2006, our investigation encompassed both the mathematics and science domains. With PISA 2006, it encompassed all three domains—reading, mathematics, and science.

In order to categorize the quality of the family background items and scales, we used the following three levels: a "weak" association existed if the variance in student achievement explained by an indicator was below 5%. A "moderate" association resulted in 5% to less than 10% of the explained variance, whereas a "strong" association was evident if 10% or more of the variance could be explained by a single indicator or a scale. This classification was again driven by our empirical findings. Our goals were, first, to compare the three studies in relation to one another, and then to identify the most appropriate or best-to-use indicator of family background within each of the three large-scale studies. This approach enabled us to set the category limits according to the empirical findings of our study.

We also computed, for the scales and indices derived in PIRLS and PISA, Pearson correlation coefficients, using the median achievement scores across countries and the respective IQRs. We determined "weak" correlations as those with coefficients of less than 0.2, "moderate" correlations as ones with coefficients ranging from 0.2 to less than 0.3, and "strong" correlations as those with coefficients equal to or greater than 0.3.

5.2.3 Reliability of Scales

Reliability is one of the most important issues in measurement. According to Cronbach (1960, p. 126), reliability provides information "about the consistency of a person's scores on a series of measurements." As DeVellis (2003, p. 27) puts it, scale reliability is "the proportion of variance attributable to the true score of the latent variable." The reliability of scales is related to the items they consist of, and also to the latent variable that these items have in common. Cronbach's alpha, widely used as a coefficient of reliability, is close to the classical definition of reliability, that is, "the proportion of variance in a scale that is attributable to the true score of the latent variable" (DeVellis, 2003, p. 47).

We used Cronbach's alpha, which indicates the internal consistency of a scale, to determine the reliability of the derived variables (scales and indices) of family background for our three studies of interest (TIMSS 2007, PIRLS 2006, PISA 2006). We consider this analysis a particularly important part of our investigation because the more reliable a scale is (in terms of the internal consistency of the variables included), the more precise is the derived measure itself. The reliability coefficient tells us how confident we can be about a measure. And because reliability is related to measurement error, the more reliable a measure is, the more precisely we can judge the influence of a certain behavior or attribute (Cronbach, 1960, p. 126).

Again, the manner in which we set criteria for describing the quality of the data was mainly driven by our empirical findings, given that our focus was on comparing results deploying data from across the three studies. Nevertheless, we discuss the outcomes of the reliability analysis with regard to commonly used quality standards (for a discussion on the latter, see Cortina, 1993). In this paper, we took a "low" reliability of a scale to mean a Cronbach's alpha of below 0.65. A "moderate" reliability meant scales or indices with a Cronbach's alpha of between 0.65 and 0.74, and a "high" reliability was achieved with a Cronbach's alpha of 0.75 and above.

To evaluate the importance of the scale items, we also investigated the item-total correlations. These statistics are used to check if the items comprising a scale measure the same construct. Correlations above 0.3 are seen as sufficient (see, for example, Everitt, 2002). Correlations below 0.3 suggest that an item should be dropped from the scale. In this paper, item-total correlations below 0.3 should be seen as "low," correlations from 0.3 to 0.5 as "moderate," and correlations above 0.5 as "high."

Another way of evaluating the importance of scale items is to investigate how the association with achievement is affected by removing or adding items. Normally, we could expect that the items that affect the scale reliability (and, hence, its validity) will

also affect the association with an outcome variable, such as academic achievement. Thus, if an item reduces or increases the reliability of a scale, it will also cause the association with the outcome variable to reduce or increase. And even if this is not the case (i. e., where there is an item that increases the scale reliability but reduces the association with achievement), the item will still be validated by the reliability analysis regardless of the association with achievement. In other words, the criterion used to validate the item is not its correlation with achievement but rather whether it is a reliable indicator of the concept at the heart of the scale. The criterion of the association with achievement is thus subordinate to the scale reliability criterion. Associations of items with an outcome variable are not theoretically meaningful on their own; what is meaningful is the association with the scale (construct), not with the individual items.

6. Findings

Using the criteria laid out in the previous chapter, we analyzed each area of family background. We report the results of each of these analyses with regard to the following issues: response-rate, association with achievement, and reliability (if applicable). The tables that we use for reporting our findings have a standardized structure:

- First, there is one table per study and area of family background and additional tables for each of the derived scales in PIRLS 2006 and PISA 2006. From the total of these 24 resulting tables, two are shown on pages 72–77 as examples. All tables are included in our online appendix to this paper as Tables A.1 to A.24. The appendix can be accessed on >http:www.ierinstitute.org/dissemination-area.html<.
- Second, all countries and benchmark participants are listed. The summary results for the participating countries (except for the benchmarking participants) are listed first. This information is followed by the median, interquartile range, and the 25th and 75th percentiles. The figures for the benchmarking participants are provided at the end of each table.
- The percentages of missing values and the percentages of explained variance (r^2) in a linear regression model are shown for all indicators for the respective area of family background and include scales, if any. The percentages of explained variance are provided for each subject area assessed by the studies of interest (i.e., mathematics, reading, and/or science).
- For the scales, reliability is provided for the full scale, followed by the item-total correlations for the individual items comprising those scales.

In order to enhance the overview of our findings, we have highlighted them according to the categories of the criteria discussed previously (i.e., "low," "moderate," and "high;" see Chapter 5.2). Table 6.1 shows the highlighting scheme of the tables in this chapter. High-quality data, regardless of the criteria, receive no highlighting. In the online appendix tables, data of a moderate quality are highlighted in yellow, whereas low-quality data are highlighted in orange.

Table 6.1: Highlighting scheme for the tables in this chapter

Highlight color	**Nonresponse (% of missing values)**	**Association with achievement (% explained variance)**	**Association with achievement (Pearson correlation coefficient)**	**Reliability of the scale (Cronbach's alpha)**	**Item-total correlation (Pearson correlation coefficient)**
[None]	Low (<5.0%)	High (>10.0%)	High (>0.3)	High (≥0.75)	High (>0.50)
[Light grey]	Moderate (5.0–9.9%)	Moderate (5.0-10.0%)	Moderate (0.20-0.30)	Moderate (0.65–0.74)	Moderate (0.30–0.50)
[Dark grey]	High (≥10.0%)	Low (<5.0%)	Low (<0.20)	Low (<0.65)	Low (<0.3)

The results from the correlation analysis are shown in separate tables in the appendix, namely Tables B.1 to B.3.

These appear on: www.ierinstitute.org/dissemination-area.html.

As examples, Table 6.2 shows the results for family's immigration status in PISA 2006 and Table 6.3 provides the results for the Early Home Literacy Activities (EHLA) scale in PIRLS 2006.

6.1 Home Possessions

(Result Tables A.1, A.2, A.3, and A.4)[1]

6.1.1 Response Rate

Information on home possessions was provided by almost all students. Throughout the age groups covered by the three studies, the median nonresponse rate was low (< 5%) in all four student populations. The items with the highest median nonresponse rates in PIRLS were the number of books in the home item (3.8%), computer possession in TIMSS Grade 4 (3.3%), and the possession of educational software in PISA (2.9%). In TIMSS Grade 8, there was no item with more than a 2% median nonresponse; here, all items showed similarly (very) low nonresponse.

The dispersion, in terms of the interquartile range (IQR) for most of the items, was low (< 5%). The items that had a moderate dispersion (IQR) across the countries were the number of books in the home in PIRLS (7.1%), the possession of a computer (6.5%), and an internet connection (5.9%) in TIMSS Grade 4.

When we took only the countries participating in both PIRLS and TIMSS Grade 4 into account, we found that Indonesia (20.3%), Kuwait (41.2%), Qatar (21.6%), and South Africa (22.4%) had high nonresponse rates for the number of books in the home in PIRLS, and that Iran (14.3%), Morocco (13.8%), Georgia (10.5%), Norway (9.6%), Germany (9.1%), and New Zealand (5.4%) showed moderate nonresponse rates.

In all these countries, the nonresponse rate was somewhat higher than the rates for the other home-possession items. Also, all of these countries, when compared to the other countries participating in both PIRLS and TIMSS Grade 4, showed a higher nonresponse rate for all home-possession items in general. In TIMSS Grade 4—although now for a different item—a similar array of countries showed higher nonresponse rates compared to the other common countries.

For the item on computer possession, Germany (14.1%), Georgia (10.8%), Morocco (9.7%), Kuwait (8.1%), Qatar (7.8%), Iran (7.3%), and Slovenia (7.3%) all had moderate nonresponse rates. Similar to the finding for PIRLS, the first five countries mentioned showed a higher nonresponse rate for all home-possession items in general when compared to the other countries.

1 Please note that all remaining tables referred to in this paper can be found in our online appendix on www.ierinstitute.org/dissemination-area.html.

Table 6.2: Example table of results for family's immigration status in PISA 2006

		Student questionnaire									
		Percent missing values					*Percent explained variance (R^2)—math achievement*				
	Country	Mother born in country (ST 11Q02)	Father born in country (ST 11Q03)	Student born in country (ST 11Q01)	Student's age of immigration (ST 11Q04)	IMMIG (Index Immigrant Background)	Mother born in country (ST 11Q02)	Father born in country (ST 11Q03)	Student born in country (ST 11Q01)	Student's age of immigration (ST 11Q04)	IMMIG
PISA 2006 participants	Argentina	1.7	2.9	1.8	25.3	2.1	0.1	0.2	0.1	3.7	0.1
	Australia	1.4	2.2	1.8	2.2	2.2	0.4	0.2	0.1	0.4	0.2
	Austria	0.8	1.2	0.6	3.3	0.8	4.4	5.3	2.5	2.3	5.1
	Azerbaijan	5.9	6.1	3.7	11.8	6.4	0.0	0.0	0.2	1.0	0.0
	Belgium	1.2	1.8	0.8	23.1	1.5	9.1	9.0	5.8	12.9	9.8
	Brazil	0.5	4.3	2.1	0.0	2.1	1.4	0.1	0.2	5.2	1.0
	Bulgaria	4.6	4.8	3.0	13.9	4.6	0.1	0.0	0.1	0.5	0.1
	Canada	4.9	5.5	4.5	1.9	5.3	0.1	0.1	0.1	0.2	0.1
	Chile	3.3	4.6	3.2	0.0	3.7	0.0	0.1	0.0	6.1	0.0
	Chinese Taipei	1.4	1.5	1.6	11.2	1.9	0.0	0.0	0.1	2.1	0.1
	Colombia	2.2	3.8	2.2	0.0	2.9	0.2	0.4	0.0	0.9	0.2
	Croatia	1.3	1.8	0.8	17.1	1.6	0.1	0.1	0.2	0.1	0.4
	Czech Republic	0.9	1.5	0.7	6.5	1.1	0.6	0.9	0.3	1.9	0.2
	Denmark	0.7	1.1	0.8	2.6	0.8	3.5	4.2	2.0	2.0	4.7
	Estonia	2.1	4.4	1.1	12.7	2.3	1.2	1.1	0.1	2.3	0.9
	Finland	0.8	1.1	0.4	4.4	0.9	1.0	1.2	1.1	2.8	1.6
	France	2.7	3.8	2.9	15.0	3.0	1.8	2.0	1.1	3.4	3.2
	Germany	5.8	6.9	4.5	14.4	5.7	6.6	6.5	2.3	1.0	5.4
	Greece	1.5	1.6	0.7	4.4	1.5	0.7	0.8	1.1	9.0	1.4
	Hong Kong SAR	0.9	1.5	1.0	6.0	1.3	0.7	0.3	2.0	5.2	1.3
	Hungary	0.9	1.3	0.7	13.8	0.9	0.0	0.1	0.0	0.4	0.0
	Iceland	2.2	2.5	1.6	3.4	2.4	0.7	0.4	0.0	0.7	1.0
	Indonesia	1.6	1.7	1.1	22.9	1.9	0.3	0.1	0.1	0.3	0.2
	Ireland	2.9	3.4	2.2	8.8	3.1	0.1	0.0	0.0	1.4	0.3
	Israel	8.7	9.1	6.5	5.8	8.4	0.1	0.2	0.1	0.0	0.0
	Italy	2.5	2.7	1.1	11.4	2.6	0.3	0.5	0.7	2.1	0.7
	Japan	0.2	1.2	0.1	2.3	0.1	0.0	0.1	0.0	4.2	0.0
	Jordan	3.8	3.8	2.2	5.4	3.9	2.3	2.4	1.1	0.8	1.4
	Korea	1.3	1.5	0.4	6.5	1.2	0.0	0.0	0.2	9.0	0.0
	Kyrgyzstan	5.6	6.4	4.2	19.2	6.1	1.1	1.8	0.1	0.1	1.0
	Latvia	2.7	5.2	1.3	4.8	2.9	0.0	0.0	0.0	3.6	0.0
	Liechtenstein	0.9	1.2	0.3	60.0	1.2	2.1	3.0	2.9	2.4	5.6
	Lithuania	2.1	3.0	1.3	17.0	2.3	0.0	0.0	0.2	0.3	0.2

Key: Moderate percentage missing value (5.0–15.0%)
Moderate percentage explained variance (5.0–9.9%)

High percentage missing value (> 15.0%)
Low percentage explained variance (< 5.0%)

— Data not available

Source: Programme for International Assessment 2006.

		Student questionnaire									
		Percent explained variance (R^2)—reading achievement					*Percent explained variance (R^2)—science achievement*				
	Country	Mother born in country (ST11Q02)	Father born in country (ST11Q03)	Student born in country (ST11Q01)	Student's age of immigration (ST11Q04)	IMMIG (Index Immigrant Background)	Mother born in country (ST11Q02)	Father born in country (ST11Q03)	Student born in country (ST11Q01)	Student's age of immigration (ST11Q04)	IMMIG
PISA 2006 participants	Argentina	0.1	0.5	0.5	2.1	0.5	0.0	0.1	0.1	3.2	0.1
	Australia	0.1	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.5	0.0
	Austria	2.4	3.0	1.1	0.8	2.7	6.6	7.9	4.7	7.4	8.7
	Azerbaijan	0.2	0.2	0.3	1.1	0.1	0.0	0.0	0.2	0.8	0.0
	Belgium	7.8	7.3	4.6	7.8	7.7	8.1	7.9	4.3	9.7	8.2
	Brazil	1.4	0.1	0.2	10.0	0.8	1.2	0.1	0.2	5.5	0.7
	Bulgaria	0.0	0.0	0.0	0.2	0.1	0.2	0.0	0.1	0.3	0.1
	Canada	0.1	0.0	0.3	1.1	0.3	0.3	0.3	0.4	0.3	0.6
	Chile	0.1	0.0	0.1	0.4	0.1	0.0	0.1	0.0	3.4	0.1
	Chinese Taipei	0.0	0.0	0.1	0.2	0.3	0.0	0.0	0.1	0.3	0.2
	Colombia	0.2	0.2	0.0	5.1	0.3	0.2	0.3	0.0	0.9	0.3
	Croatia	0.1	0.2	0.1	0.0	0.3	0.2	0.3	0.3	0.4	0.6
	Czech Republic	0.9	0.9	0.9	7.9	0.7	0.7	0.9	0.5	3.2	0.4
	Denmark	2.9	3.8	1.8	3.3	4.2	4.3	5.1	1.9	1.0	5.5
	Estonia	4.4	4.1	0.9	0.1	3.8	2.0	1.8	0.1	2.3	1.6
	Finland	0.4	0.6	0.7	1.7	0.8	0.8	1.2	1.8	8.2	1.9
	France	0.7	1.0	0.6	1.1	1.5	1.6	1.9	1.2	2.1	3.1
	Germany	6.0	5.6	2.0	1.3	4.9	9.1	8.7	3.0	2.0	7.4
	Greece	0.5	0.4	0.7	3.2	0.8	1.0	1.0	1.5	9.9	1.8
	Hong Kong SAR	0.3	0.1	1.8	3.3	0.7	0.4	0.1	1.6	4.9	0.8
	Hungary	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0
	Iceland	0.7	0.6	0.2	4.4	1.4	0.8	0.5	0.0	2.3	1.2
	Indonesia	0.5	0.4	0.3	0.6	0.5	0.4	0.2	0.2	0.5	0.3
	Ireland	0.1	0.0	0.0	1.0	0.1	0.0	0.0	0.0	0.7	0.1
	Israel	0.1	0.1	0.0	0.9	0.0	0.0	0.0	0.1	0.3	0.0
	Italy	0.5	0.9	1.1	4.9	1.2	0.6	1.0	1.1	5.8	1.4
	Japan	0.1	0.1	0.0	1.7	0.1	0.0	0.1	0.0	5.2	0.1
	Jordan	2.4	2.7	1.4	0.5	1.4	2.0	2.5	0.9	0.3	1.2
	Korea	0.1	0.0	0.1	7.6	0.1	0.0	0.0	0.1	21.7	0.0
	Kyrgyzstan	0.2	1.0	0.0	0.0	0.3	0.8	1.6	0.1	0.1	0.6
	Latvia	0.1	0.1	0.1	1.4	0.3	0.0	0.0	0.0	4.8	0.0
	Liechtenstein	3.2	4.2	5.5	4.8	8.4	2.4	3.4	2.6	2.6	6.2
	Lithuania	0.1	0.2	0.1	1.3	0.0	0.0	0.1	0.1	0.7	0.0

Table 6.2: Example table of results for family's immigration status in PISA 2006 (contd.)

		Student questionnaire									
		Percent missing values					*Percent explained variance (R^2)—math achievement*				
	Country	Mother born in country (ST 11Q02)	Father born in country (ST 11Q03)	Student born in country (ST 11Q01)	Student's age of immigration (ST 11Q04)	IMMIG (Index Immigrant Background)	Mother born in country (ST 11Q02)	Father born in country (ST 11Q03)	Student born in country (ST 11Q01)	Student's age of immigration (ST 11Q04)	IMMIG
PISA 2006 participants	Luxembourg	1.4	2.2	1.0	6.6	1.7	5.2	6.6	2.9	0.9	6.3
	Macao-China	1.1	1.5	1.6	3.4	1.9	0.2	0.4	0.3	0.5	0.0
	Mexico	2.6	4.1	3.8	28.4	4.6	2.2	2.3	1.7	0.0	2.6
	Montenegro	2.8	3.1	2.4	7.7	3.2	0.6	0.8	1.0	0.6	0.2
	Netherlands	1.3	1.5	1.4	6.0	1.9	4.5	4.5	0.9	4.9	4.3
	New Zealand	2.1	2.8	1.9	1.7	2.4	0.1	0.0	0.1	0.6	0.0
	Norway	2.3	2.4	2.1	4.7	2.4	0.7	1.4	1.0	0.5	1.9
	Poland	1.7	2.1	0.7	6.2	1.8	0.0	0.0	0.0	5.9	0.0
	Portugal	0.9	1.4	1.1	4.6	1.2	0.6	0.9	0.9	0.0	1.8
	Qatar	8.6	8.6	6.0	9.5	8.7	8.4	9.7	8.8	11.1	11.7
	Romania	0.1	0.3	0.0	49.7	0.1	0.0	0.0	0.0	25.7	0.0
	Russian Federation	1.3	3.5	0.8	7.7	1.4	0.1	0.1	0.0	1.5	0.3
	Serbia	1.7	2.1	1.0	8.7	1.9	0.3	0.6	0.1	0.0	0.2
	Slovak Republic	1.0	1.5	0.8	10.4	1.1	0.2	0.0	0.1	0.2	0.1
	Slovenia	1.1	1.7	0.8	6.2	1.3	1.2	1.4	0.2	1.9	1.4
	Spain	1.2	1.9	1.0	1.1	1.5	1.8	1.7	2.3	1.2	2.7
	Sweden	1.3	1.6	1.6	2.7	1.7	2.2	3.5	2.1	0.7	3.4
	Switzerland	1.2	1.2	0.9	6.7	1.4	7.5	7.7	6.3	0.6	10.3
	Thailand	2.9	3.0	1.5	22.6	3.3	0.0	0.1	0.0	6.1	0.1
	Tunisia	2.0	1.6	1.5	47.7	1.7	0.0	0.3	0.0	15.1	0.5
	Turkey	2.3	2.3	0.8	3.9	2.5	0.3	0.1	0.1	2.9	0.2
	United Kingdom	2.5	3.5	2.8	4.4	3.4	0.2	0.5	0.2	0.1	0.6
	United States	2.4	3.6	2.1	8.5	3.2	1.1	1.1	1.0	0.1	1.4
	Uruguay	3.4	4.5	3.6	11.5	4.3	0.1	0.0	0.1	0.4	0.0
	Median	1.7	2.3	1.4	6.6	2.1	0.3	0.4	0.2	1.2	0.4
	Interquartile range	1.5	2.2	1.4	9.4	1.9	1.4	1.7	1.0	3.2	1.7
	25th percentile	1.2	1.5	0.8	4.4	1.4	0.1	0.1	0.1	0.4	0.1
	75th percentile	2.7	3.8	2.2	13.8	3.2	1.4	1.7	1.1	3.6	1.8

Key: Moderate percentage missing value (5.0–15.0%)
Moderate percentage explained variance (5.0–9.9%)

High percentage missing value (> 15.0%)
Low percentage explained variance (< 5.0%)

— Data not available

Source: Programme for International Student Assessment 2006.

		Student questionnaire									
		Percent explained variance (R^2)—reading achievement					*Percent explained variance (R^2)—science achievement*				
	Country	Mother born in country (ST11Q02)	Father born in country (ST11Q03)	Student born in country (ST11Q01)	Student's age of immigration (ST11Q04)	IMMIG (Index Immigrant Background)	Mother born in country (ST11Q02)	Father born in country (ST11Q03)	Student born in country (ST11Q01)	Student's age of immigration (ST11Q04)	IMMIG
PISA 2006 participants	Luxembourg	7.6	8.9	3.2	1.8	8.7	8.6	10.0	3.2	1.3	9.4
	Macao-China	0.4	0.9	0.1	0.0	0.2	0.2	0.6	0.8	0.5	0.0
	Mexico	3.2	3.0	2.0	0.1	3.5	2.9	2.7	2.3	0.1	3.6
	Montenegro	0.7	0.9	0.9	0.9	0.3	0.7	0.9	1.4	0.8	0.4
	Netherlands	3.5	3.4	1.2	8.2	3.8	5.2	4.9	1.1	8.0	5.2
	New Zealand	0.1	0.2	0.2	1.6	0.4	0.1	0.2	0.0	1.2	0.2
	Norway	0.6	1.2	1.1	0.9	1.6	0.9	1.5	1.1	1.1	2.0
	Poland	0.0	0.0	0.0	5.7	0.1	0.0	0.0	0.0	3.9	0.0
	Portugal	0.6	0.8	1.0	0.0	1.9	0.8	1.0	1.0	0.0	2.3
	Qatar	8.5	9.2	9.6	10.4	12.1	10.2	11.6	11.0	12.5	14.7
	Romania	0.0	0.1	0.7	30.9	0.0	0.1	0.2	0.2	44.9	0.1
	Russian Federation	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	1.2	0.2
	Serbia	0.3	0.5	0.1	0.2	0.1	0.3	0.4	0.1	0.2	0.1
	Slovak Republic	0.1	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.2	0.1
	Slovenia	0.8	1.0	0.1	1.1	0.9	2.3	2.5	0.4	6.2	2.6
	Spain	1.3	1.3	1.9	0.1	2.2	1.7	1.7	2.5	0.5	3.0
	Sweden	1.6	1.7	1.9	4.9	2.6	2.7	3.8	2.6	1.6	4.4
	Switzerland	5.8	6.3	6.1	0.1	9.3	8.4	8.8	6.8	0.2	11.7
	Thailand	0.0	0.1	0.0	3.6	0.0	0.1	0.3	0.0	16.6	0.3
	Tunisia	0.1	0.5	0.0	13.3	0.7	0.0	0.4	0.0	5.6	0.6
	Turkey	0.0	0.0	0.0	0.2	0.0	0.1	0.1	0.0	0.2	0.1
	United Kingdom	0.1	0.3	0.4	0.1	0.6	0.2	0.5	0.3	0.1	0.8
	United States	–	–	–	–	–	2.2	2.3	1.5	0.0	2.7
	Uruguay	0.1	0.0	0.1	0.0	0.0	0.1	0.0	0.2	0.5	0.0
	Median	0.3	0.5	0.3	1.1	0.5	0.4	0.4	0.2	1.2	0.6
	Interquartile range	11.3	1.1	1.0	3.6	1.5	1.9	1.8	1.4	4.6	2.5
	25th percentile	0.1	0.1	0.1	0.2	0.1	0.0	0.1	0.1	0.3	0.1
	75th percentile	1.3	1.2	1.1	3.8	1.7	2.0	1.9	1.5	4.9	2.6

Table 6.3: Example table of results for the Early Home Literacy Activities (EHLA) scale in PIRLS 2006

					Item-total correlation					
	Country	Percent missing values	Percent explained variance (r^2)–reading achievement	Full scale reliabilities–Cronbach's alpha	*Reading books (ASBHHA01)*	*Telling stories (ASBHHA02)*	*Singing songs (ASBHHA03)*	*Playing with ABC-tools (ASBHHA04)*	*Playing word games (ASBHHA07)*	*Reading aloud (ASBHHA09)*
PIRLS 2006 participants	Austria	5.6	5.0	0.64	0.39	0.41	0.33	0.33	0.44	0.36
	Belgium (Flemish)	3.9	4.1	0.67	0.39	0.38	0.34	0.38	0.49	0.39
	Belgium (French)	11.5	3.7	0.64	0.39	0.41	0.26	0.36	0.46	0.33
	Bulgaria	4.6	3.4	0.79	0.62	0.60	0.36	0.57	0.58	0.52
	Chinese Taipei	6.0	5.7	0.74	0.51	0.53	0.39	0.48	0.48	0.50
	Denmark	6.5	2.4	0.66	0.41	0.40	0.36	0.34	0.49	0.33
	England	53.9	3.9	0.72	0.45	0.42	0.44	0.42	0.53	0.46
	France	9.4	2.8	0.63	0.39	0.39	0.28	0.38	0.43	0.33
	Georgia	8.8	1.6	0.70	0.50	0.35	0.35	0.50	0.49	0.46
	Germany	13.8	1.3	0.61	0.30	0.37	0.32	0.32	0.43	0.34
	Hong Kong SAR	5.4	1.6	0.73	0.53	0.55	0.44	0.39	0.51	0.41
	Hungary	10.7	2.3	0.63	0.38	0.38	0.32	0.39	0.42	0.29
	Iceland	24.3	3.4	0.69	0.38	0.45	0.36	0.42	0.52	0.40
	Indonesia	4.6	2.7	0.73	0.47	0.43	0.41	0.55	0.54	0.43
	Iran, Islamic Republic of	4.5	7.2	0.74	0.51	0.45	0.44	0.47	0.55	0.47
	Israel	38.1	0.1	0.70	0.40	0.47	0.36	0.49	0.48	0.42
	Italy	6.9	1.9	0.60	0.27	0.37	0.27	0.36	0.41	0.35
	Kuwait	28.0	2.3	0.66	0.45	0.41	0.16	0.47	0.49	0.40
	Latvia	6.8	2.2	0.62	0.39	0.40	0.30	0.36	0.40	0.28
	Lithuania	4.3	1.9	0.64	0.39	0.39	0.29	0.41	0.43	0.33
	Luxembourg	8.6	4.6	0.69	0.45	0.44	0.34	0.42	0.50	0.40
	Macedonia	8.4	2.3	0.69	0.43	0.43	0.38	0.46	0.45	0.41
	Moldova	3.5	3.5	0.69	0.51	0.48	0.32	0.47	0.36	0.42
	Morocco	4.8	3.3	0.73	0.53	0.46	0.35	0.51	0.44	0.53
	Netherlands	33.2	1.8	0.66	0.30	0.35	0.34	0.42	0.50	0.44
	New Zealand	35.5	5.5	0.77	0.52	0.53	0.44	0.49	0.58	0.51
	Norway	7.6	1.7	0.65	0.35	0.38	0.32	0.42	0.44	0.37
	Poland	3.6	3.4	0.63	0.37	0.36	0.30	0.42	0.41	0.34
	Qatar	33.0	1.4	0.62	0.34	0.42	0.16	0.41	0.44	0.37
	Romania	3.6	13.6	0.78	0.56	0.59	0.39	0.58	0.58	0.48
	Russian Federation	1.8	4.0	0.68	0.47	0.41	0.33	0.44	0.47	0.36
	Scotland	48.4	5.4	0.72	0.46	0.45	0.40	0.45	0.53	0.48
	Singapore	2.8	4.3	0.79	0.58	0.59	0.49	0.54	0.52	0.56
	Slovak Republic	3.4	2.5	0.61	0.36	0.37	0.24	0.35	0.42	0.34
	Slovenia	6.0	2.6	0.67	0.38	0.43	0.36	0.38	0.48	0.39
	South Africa	16.1	2.2	0.59	0.31	0.27	0.22	0.38	0.39	0.39
	Spain	38.5	3.6	0.64	0.35	0.42	0.36	0.39	0.42	0.28
	Sweden	6.7	1.9	0.69	0.36	0.40	0.38	0.46	0.53	0.39
	Trinidad and Tobago	13.2	6.8	0.73	0.52	0.52	0.39	0.45	0.49	0.45
	United States	–	–	–	–	–	–	–	–	–

	Country	Percent missing values	Percent explained variance (r^2)—reading achievement	Full scale reliabilities—Cronbach's alpha	Item-total correlation					
					Reading books (ASBHHA01)	*Telling stories (ASBHHA02)*	*Singing songs (ASBHHA03)*	*Playing with ABC-tools (ASBHHA04)*	*Playing word games (ASBHHA07)*	*Reading aloud (ASBHHA09)*
	Median	6.9	2.8	0.68	0.40	0.42	0.35	0.42	0.48	0.40
	Interquartile range (IQR)	10.4	2.0	0.09	0.13	0.07	0.07	0.09	0.08	0.11
	25th percentile	4.6	2.0	0.64	0.37	0.39	0.31	0.38	0.43	0.34
	75th percentile	14.9	4.0	0.73	0.50	0.46	0.38	0.47	0.51	0.45
PIRLS 2006 benchmark participants	Canada (Alberta)	20.6	4.1	0.74	0.47	0.48	0.43	0.49	0.53	0.49
	Canada (British Columbia)	23.9	2.7	0.75	0.46	0.48	0.49	0.45	0.52	0.53
	Canada (Nova Scotia)	9.8	3.0	0.73	0.46	0.48	0.40	0.45	0.52	0.51
	Canada (Ontario)	11.4	2.2	0.73	0.47	0.49	0.42	0.46	0.48	0.48
	Canada (Québec)	10.4	2.8	0.68	0.45	0.45	0.34	0.43	0.45	0.38
	Iceland (Grade 5)	33.9	4.9	0.70	0.43	0.45	0.36	0.45	0.51	0.41
	Norway (Grade 5)	10.2	2.5	0.64	0.33	0.38	0.26	0.42	0.46	0.37

Key:

Moderate percentage missing value (5.0–15.0%)
Moderate percentage explained variance (5.0–9.9%)
Moderate reliability (0.65–0.74)
Moderate item-total correlation (0.30–0.50)

High percentage missing value (> 15.0%)
Low percentage explained variance (< 5.0%)
Low reliability (< 0.65)
Low item-total correlation (< 0.30)

— Data not available

Source: Progress in International Reading Literacy Study 2006.

For the TIMSS Grade 8 and PISA common countries, almost all items showed low nonresponse with respect to the TIMSS Grade 8 data. Only Romania (6.1%) and Tunisia (5.9%) had moderate nonresponse rates for the computer-possession item. For the same set of countries, the PISA data consisted of slightly more nonresponse. Whereas Qatar and Israel had moderate nonresponse rates for all (19) or almost all (17) international home-possession items, Jordan (10) and Tunisia (7) showed moderate nonresponse for at least several items, and Colombia (3) and Romania (1) showed moderate nonresponse for fewer items.

PIRLS also collected (from parents through the home questionnaire) possession information on the number of books as well as children's books in the home. In both cases, the median nonresponse rate was moderate, with 6.3% and 6.1%, respectively. The dispersion (IQR) in both cases was high (11.5% and 11.7%). However, unlike the outcome with respect to the student data, we could find no particular issues with either one of the two variables (which are the only home-possession items included in the PIRLS 2006 home questionnaire). Instead, the country-level results suggested that the high nonresponse rates in several countries were due to unit nonresponse. Moreover, the two items on home possessions showed the least amount of nonresponse in almost all countries, compared to all other items on family background included in the PIRLS home questionnaire.

6.1.2 Association with Achievement

Our investigation into the association of indicators of home possessions with achievement showed that, compared to the other home-possession single items, information on number of books in the home offered the highest amount of explained variance in all three subject domains (reading, mathematics, and science) and in all four student populations. With PISA, we found a strong association between books in the home and student achievement, with a median of the explained variance of 12.6% in science, 12.2% in mathematics, and 11.1% in reading (IQRs between 5.2% and 7.1%). For TIMSS Grade 8, we found a moderate association (8.0% and 8.1%) for both mathematics and science achievement (IQRs of 7.8% and 10.2%). In PIRLS, the median of the variance in reading achievement explained by the number of books in the home (as reported by students) was also moderate, with 8.6% (an IQR of 5.4%). The number of books in the student's home, as reported by the parents, as well as the number of children's books in the home again had a moderate association with students' achievement, with a median of 8.5% and 9.3% of the explained variance (IQRs of 4.7% and 6.7%, respectively), an outcome that is quite similar to the information on books in the home provided by students. Although the medians of the explained variances in mathematics and science achievement in TIMSS Grade 4 were quite similar at 5.5% and 6.3%, respectively (and with IQRs of 7.2% and 6.9%), they were still moderate.

Another variable associated with book owning has been used only in PISA. Possession of classic literature showed a moderate association with achievement, with a median explained variance of 6.1% in mathematics, 6.9% in reading, and 7.1% in science (IQRs of 3.5%, 5.6%, and 4.2%, respectively). The possession of poetry (again used in

PISA only) showed a much weaker relationship, with achievement in all three subject domains. In PISA, an additional item showed moderate associations with student achievement. The number of cars at home accounted for 6.4% of the variance in mathematics, 5.4% in science, and 3.9% in reading achievement (IQRs of 5.1%, 4.6%, and 4.4%). These items are not surveyed in TIMSS and PIRLS.

Students possessing their own books at home was the second-best home-possession single indicator of family background for reading achievement in PIRLS, with a median-explained variance in a simple linear regression model of 4.5% (IQR of 3.1%), which was low according to our classification, and also lower than the variables associated with number of books. No items amongst the TIMSS Grade 4 home possession items were similarly outstanding.

Within the older age group (TIMSS Grade 8 and PISA), possessing a computer had some explanatory power in several countries. However, on average, the median-explained variance was low for both studies and all subject domains. In PISA, the median-explained variances ranged from 3.9% to 4.5% in the three subject domains (IQRs from 4.6% to 7.6%). In TIMSS Grade 8, the median-explained variance in mathematics achievement was 3.1% (IQR of 4.2%), and the median-explained variance in science achievement was 2.2% (IQR of 3.6%).

Another possession item closely related to the possession of a computer at home is the availability of an internet connection. This item was included in TIMSS (both grades) and in PISA. For the older age group, results were similar to those for computer possession. Although availability of internet access at home had some explanatory power in several countries, on average, across all countries, it showed only a low association with achievement. In PISA, the median explained variance was 4.3% for mathematics (IQR of 4.7%), 3.7% for science achievement (IQR of 4.7%), and 3.3% for reading achievement (IQR of 4.0%). In TIMSS Grade 8, it was 2.8% for mathematics and 2.5% for science achievement (IRQs of 4.1% and 3.5%). In TIMSS Grade 4, this variable yielded a median of only 1.9% of the explained variance in mathematics and 1.7% in science achievement (IRQs of 4.1% and 3.5%), suggesting that the internet may not be widely used by students in the younger age group.

For the rest of the variables related to home-possession items, the associations with achievement in either subject domain were low, with medians of explained variance below 5.0%. Because of this result, we consider there is no need to discuss it in detail here.

Our analysis of home-possession items suggests that the possession of books of whatever type was the variable most strongly associated with academic achievement across all four student populations and their combinations of different home-possession items. This kind of possession item is used in many studies as a proxy for cultural capital; some researchers even use it as the only proxy for SES. Regardless of whether using only possession of books at home to represent SES is reasonable or not, the different variables related to availability of books at home all showed a moderate median association with achievement.

The second strongest association between achievement and the possession items used in all studies lay with the variables related to availability of computer access and internet connection at home, although their relationship with academic achievement was much weaker. All other home-possession items, when used as single predictors in the three studies, showed negligible relationships with educational outcomes in terms of achievement

When we look at country-level results, we can see that the association of possession of books with students' academic achievement was quite diverse across countries. In PIRLS, the association ranged from almost none (0.4%) in Qatar to more than 17.0% in the Slovak Republic, Luxembourg, and Hungary. The TIMSS Grade 4 data showed a similar range, from almost no association, for example, with mathematics achievement in Armenia and Kazakhstan, to 16.6% in Hungary. The TIMSS Grade 4 and PIRLS common countries showed a similar pattern, somewhat independent from the subject area and the study. In Hong Kong SAR, Italy, Kuwait, Morocco, and Qatar, for example, the association between book possession (as reported by students) and achievement was low in both studies and subject areas. In turn, the association was high for another common set of countries (Austria, Chinese Taipei, England, Germany, Hungary, and New Zealand). Similar results emerged from the data for the older age group. In summary, although the association of book possession with achievement was diverse across countries, there were also similar patterns. The same three countries (of the common country sets) had a low association (Indonesia, Jordan, and Qatar), again regardless of the subject area assessed.

6.2 Immigration Status

(Result Tables A.5, A.6, A.7, and A.8)

6.2.1 Response Rate

Immigration status was operationalized in the studies we considered to include three aspects—the immigration status of the parents, the immigration status of the student, and the age of the student when he or she immigrated (the latter information was not collected in PIRLS).

The amount of nonresponse for the variables associated with immigration status was low, in general, for TIMSS and PISA. The age at which students immigrated revealed higher nonresponse than the information on whether any of the three family members (mother, father, student) was born in the country of the test. Whereas in TIMSS Grade 8, the median nonresponse rate for information on students' age of immigration was still low, at 4.5% and an IQR of 4.4%, PISA had a moderate median nonresponse rate of 6.6% (IQR of 9.4%). This result might be due to the different item formats. PISA required students to enter a number; TIMSS used a closed-item format with three categories.

A look at country level in the older age group shows only a few countries in TIMSS Grade 8 and PISA with moderate nonresponse rates for the information on parents' and students' country of birth. In the main, nonresponse for the immigration-age information varied a lot across the PISA countries (from almost none in Brazil, Chile,

and Colombia to 60.0% in Lichtenstein). The variation in TIMSS Grade 8 was much smaller: from almost none in Italy and Korea to 32.7% in Iran, which was a single outlier.

Item format may also be the reason for the moderate nonresponse rates in PIRLS with respect to the information on whether the student's mother and father were born in the country of the test. PIRLS offered students a "don't know" option. When combined with the percentages of omitted responses, the median missing rate was 5.3% (IQR of 6.9%) for the mother's and 6.3% (IQR of 9.1%) for the father's immigration status. Also, the indicator combining both parents' immigration status had a moderate nonresponse rate, with a median of 5.1% (IQR of 6.9%) across countries. However, the information on whether the student was born in the country of the test showed a low nonresponse rate of 2.4% (IQR of 5.7%).

In comparison with the TIMSS Grade 4 data, the data pertaining to the PIRLS countries varied considerably more in relation to nonresponse for information about the country of birth of parents and students. In PIRLS, several countries had a high nonresponse rate for immigrant-status information. Kuwait, for example, showed a percentage of 34.4 for mother's immigration status, whereas in TIMSS Grade 4 only a few countries showed even moderate nonresponse rates (with a maximum of 14.0% for father's immigration status in Germany).

6.2.2 Association with Achievement

Most of the variables related to the immigration status of the students and their parents showed only very weak associations, on average, with all subject domains in terms of explained variance in a linear regression model. With regard to the younger age group, no indicator in PIRLS showed a reasonable association with achievement across all countries (the maximum median-explained variance was 1.0%). The regression model with the combined parents' immigration status was not even significant for the majority of the PIRLS countries in our focus. In TIMSS Grade 4, the combined parents' immigration status also showed only a weak association, with a median of 1.1% (IQR of 1.8%) of the explained variance in mathematics achievement and 1.3% (IQR of 2.7%) in science achievement. Students' immigration status yielded a median of 3.2% (IQR of 5.2%) of the explained variance in mathematics achievement and 2.6% (IQR of 5.8%) in science achievement. In similar vein, only three of the PIRLS countries (Austria, Germany, and Luxembourg) had moderate or strong associations with the separate sets of information on mother's and father's immigration status and student achievement, whereas in TIMSS Grade 4, a good many more countries showed moderate and strong associations. This outcome was especially the case for students' immigration status, which was moderately (13 countries for mathematics and 12 countries for science achievement) or even strongly (two countries for mathematics and three countries for science achievement) associated with students' achievement.

It is interesting, though, that the students' immigration status showed a higher association in the TIMSS Grade 4 than in the TIMSS Grade 8 data. In the older age group, almost no variance was explained by either of the indicators of immigration status (median maximum of 2.1% for TIMSS Grade 8 and 1.2% for PISA). In PISA, not one significant relationship emerged in about half of the countries; and of the TIMSS Grade 8 countries, only a few countries showed such a significant relationship.

A more detailed look at the results at the country level revealed no clear pattern regarding the countries common to TIMSS Grade 8 and PISA. Here, the association between a student's or a family's immigration status was only moderately correlated with the student's achievement. In TIMSS Grade 8, Chinese Taipei (6.4% in mathematics and 6.5% in science achievement) and Indonesia (5.8% in mathematics and 7.1% in science achievement) showed a moderate association between student achievement and whether or not the student was born in the country of the test.

The age at which a student immigrated was moderately associated with the student's achievement in several countries, but the sets of countries differed according to subject area. For mathematics achievement, Japan, Korea, and Romania showed higher associations than did the other countries of interest. For science achievement, this was the case in Israel, Japan, Norway, the Russian Federation, and Thailand. In PISA 2006, Colombia, the Czech Republic, Italy, Japan, Korea, Qatar, Romania, Slovenia, Thailand, and Tunisia all showed a moderate association between some of the immigration variables and students' achievement. When we used data available from the United Nations (2010) to categorize countries according to their immigration policies, we found that the picture regarding the above findings was no clearer.

The set of common countries in the younger age-group studies (PIRLS and TIMSS Grade 4) was quite similar in terms of the immigration status of the parents, with Austria, Germany, and the Netherlands showing moderate associations with achievement. However, when we looked at just the status of the student, both studies yielded very different results. In PIRLS, this indicator seemed to have no explanatory power at all, whereas in TIMSS Grade 4 almost half of the countries in scope showed moderate or even strong associations.

6.3 Language Used at Home

(Result Tables A.9, A.10, A.11, and A.12)

6.3.1 Response Rate

Information that participating students provided on the language used at home had a low level of nonresponse in TIMSS and PISA, with a median of 1.0% in TIMSS Grade 4, 0.9% in TIMSS Grade 8, and 2.7% in PISA (IQRs of 2.7%, 1.1%, and 2.6%). In PIRLS, the median nonresponse level was moderate at 5.8% and an IQR of 7.8%. PIRLS was the only one of the three large-scale studies that asked parents as well as students about the language spoken at home. The median amount of nonresponse was high at 15.6% for language used by the father, and it was moderate (but close to high at 14.7%) for the language used by the mother (IQRs of 15.6% and 16.5%, respectively).

The information on the language of children's books that parents provided was somewhat more complete, but still on a moderate level, with a median of 7.6% of missing data (IQR of 12.3%). More indepth analysis suggested that the large amount of missing data provided by parents compared to data provided by students was due to parents' nonparticipation in the survey. Also, some parents might have had no interest in completing the questionnaire or answering all questions. Another reason could be that only one adult (parent, guardian, or the like) answered the questionnaire, and that he or she was not willing or able to provide the information for a second adult (parent) living with him or her and the child.

Among the common countries of the younger age group, Germany, Kuwait, and Morocco all had a moderate or high nonresponse rate in both TIMSS Grade 4 and PIRLS for student language use. In PIRLS, several more countries showed a moderate or high nonresponse rate, such that no clear pattern emerged. Nevertheless, with response rates in PIRLS ranging from 1.3% in Luxembourg to 33.2% in Kuwait, the amount of missing data—and thus information on the language used at home—seems to have differed to quite a strong degree across countries. Data for the older age group were, in general, more complete. The Grade 8 data for the countries held in common by PISA and TIMSS 2007 showed a nonresponse rate of less than 2.0% in each country. In PISA, only Italy, Israel, Qatar, and Tunisia had moderate nonresponse rates; all other common countries had low nonresponse rates. A few of the PISA countries that did not participate in TIMSS also showed moderate nonresponse rates.

6.3.2 Association with Achievement

Our analysis showed almost no association, in PIRLS, between the language spoken at home, as reported by the students, and students' reading achievement (the median-explained variance was 0.8% and the IQR was 1.3%). We found similar results for the language parents used when talking to the student (median of 0.8% with an IQR of 1.9% for fathers, and 0.9% with an IQR of 1.4% for mothers). The median association between the language of children's books and students' achievement was even lower (0.2% with an IQR of 0.5%). In PIRLS, only three countries showed a moderate (Austria and Luxembourg) or high (Iran) association between language spoken at home, again as reported by students, with reading achievement. The information on the language that parents used at home yielded moderate or high associations in just a few more countries. Four countries showed a moderate (Austria, Bulgaria, and Singapore) or high (Iran) association between mothers' language use and reading achievement. Four countries (Bulgaria, Chinese Taipei, Germany, and Scotland) showed a moderate association between fathers' language use and student achievement, and one (Iran) showed a high association.

Analysis of the TIMSS Grade 4 data revealed a low association across countries for both mathematics achievement (1.5% and an IQR of 2.1%) and science achievement (2.0% and an IQR of 3.5%). Four TIMSS Grade 4 countries (Chinese Taipei, Hong Kong SAR, Iran, and Singapore) showed a moderate association with mathematics achievement, and eight countries showed a moderate (Austria, Chinese Taipei, Germany, Hong

Kong SAR, Iran, and the United States) or a high (Qatar and Singapore) association with science achievement.

Language use at home was operationalized differently in PISA and TIMSS Grade 8. TIMSS asked students to report, using four response categories ("always," "almost always," "sometimes," and "never"), on the frequency with which the language of the test was spoken at home. PISA simply asked if the most frequent language used at home was the language of the test. Regardless of how the question was formulated, the items in both studies did not explain a substantive amount of variance across countries (median maximum of 0.9%), so categorizing the associations as weak. In PISA, we found high or moderate associations only in Liechtenstein (high), Bulgaria, Germany, and Switzerland (moderate) with all three subject areas assessed. In Austria, we found a moderate association with mathematics and science achievement, and in Chinese Taipei the same level of association with reading achievement. In TIMSS Grade 8, four countries (Chinese Taipei, Hong Kong SAR, Iran, and Turkey) showed a moderate association with mathematics and science achievement. Another three countries showed a moderate (Singapore and Thailand) or a high (Qatar) association with science achievement only.

It should be noted that most of the variables associated with the use of different languages are at a nominal level of measurement, a situation that poses limitations on the types of analyses that researchers can perform. More specifically, the only correlation or regression analysis that can be performed with these analyses to test the relationship with achievement is simple linear regression, in which the language of interest is coded as "1" and all the rest of the values are coded as "0." This type of coding is called "dummy coding," and we applied it, where relevant, for our analysis purposes.

6.4 Parental Education

(Result Tables A.13, A.14, and A.15)

Earlier in this paper, we delineated the role that parental education, one of the constituents of the family background construct, plays with respect to students' achievement in school. Our analysis in this paper supported research literature findings that parents' education is a family characteristic which, in general, has an association with achievement. In the case of our analysis, the association was a moderate one.

The two large-scale studies conducted at the Grade 4 level did not ask the participating students to provide information about their parents' education. Thus, the TIMSS Grade 4 data contain no information on this variable. In PIRLS, this information was collected via the home questionnaire completed by students' parents. In PISA (involving the older age group), both students and their parents were asked to provide this information. TIMSS Grade 8 collected these data from students only.

6.4.1 Response Rate

In general, the median amount of nonresponse was again much higher in the data provided by parents (applicable for PISA and PIRLS). The 20.2% median nonresponse for the information on father's education and 17.6% for mother's education in PISA (IQRs of 25.5% and 26.8%) along with the 15.9% median nonresponse for father's education and 16.1% for mother's education in PIRLS (IQRs of 18.5% and 20.1%) led to us classifying all items at the high nonresponse level. Although the combined measure of the highest education of either parent had less missing data, the nonresponse rate was nevertheless moderate—13.3% in PIRLS (IQR of 18.3%) and 15.0% in PISA (IQR of 24.6%). As the high IQRs already indicate, nonresponse rates differed considerably across countries. For example, whereas the nonresponse rate for all items on parental education was low in the PIRLS data for the Russian Federation (3.6%, 2.9%, and 2.9%), the rates were extremely high in England (57.2%, 57.1%, and 56.2%). A similar diversity was apparent in the PISA home questionnaire data.

The student data nonresponse rates on parental education in PISA were low in absolute terms and also much lower compared to parental nonresponse rates. The median missing rates ranged from 1.1% (IQR of 2.2%) for the combined indicator of parental education to 4.5% (IQR of 4.4%) for father's education. Accordingly, the differences across countries were much smaller. Only a few countries showed moderate nonresponse rates for information on the mother's or the father's education. In all other countries, the rate was low.

The TIMSS Grade 8 student-provided data were exceptional because a "don't know" option was again offered to students when they were asked to report their parents' educational attainment. The percentage of students who chose the don't know option and those who omitted the item added up to a moderate median nonresponse rate for the mother's education and the father's education of 8.7% and 10.5% (IQRs of 11.8% and 19.2%), respectively. The combined information about the parents' highest education level yielded a lower—but still moderate—level of missing data (5.9% with an IQR of 9.4%). Again, the countries' nonresponse rates differed to quite a marked degree. For example, on the one hand, Malaysia showed very low rates (0.5% for mother's education, 0.8% for father's education, and 0.2% for the highest of both), while on the other hand, Sweden had extremely high nonresponse rates (58.5%, 58.2%, and 52.0%, respectively).

6.4.2 Association with Achievement

The variables on parental education showed, in most cases, moderate associations with achievement, a pattern that held across all the studies, age groups, and respondent levels (students and parents). The medians of explained variance in achievement ranged from 5.4% to 11.9% (IQRs of 2.7% to 6.7%). In PIRLS, mother's education and the parents' highest education yielded a strong association with reading achievement of 10.7% and 11.9% (IQRs of 5.7% and 6.1%), respectively. The only low association with students' reading achievement, with 4.6% of the explained variance (IQR of 3.8%), was found for the student-reported information in PISA about mother's education.

The highest amounts of explained variance for variables related to education, as reported by parents in PISA, were for mathematics achievement, where the explained variance was 6.7% for the mother, 7.1% for the father, and 8.5% for the highest parental education (IQRs of 3.4%, 4.6%, and 3.8%, respectively). The amount of explained variance in achievement in PISA for the education variables (as reported by the students) ranged from 4.6% to 5.4% in reading, 5.8% to 6.7% in science, and 6.1% to 7.1% in mathematics (IQRs of 3.8% to 4.7%, 4.8% to 5.4%, and 4.5% to 5.6%). In PISA, the educational variables, again as reported by the students, showed the strongest association with mathematics achievement. However, the differences with respect to the other cognitive domains were relatively small. When compared to the student-provided data, the information provided by the parents again tended to show slightly stronger associations with achievement. Both the information supplied by students and that supplied by parents on parental education showed somewhat diverse associations across countries, with all three levels of association (weak, moderate, and strong) evident with respect to students' achievement.

The PIRLS data on parental education, collected solely from parents, yielded moderate to strong associations with students' achievement, and were thus the strongest across all three studies and cognitive domains (9.8% for father's education, 10.7% for mother's education, and 11.9% for the highest education of both). Countries differed only to a small extent. In nearly all countries, the associations were moderate or even strong; the only exceptions were the three Canadian provinces, Hong Kong SAR, Italy, and Qatar.

The TIMSS Grade 8 data consisted of variables on parental education obtained from students only. Compared to the association between parental education and student achievement in PISA, the TIMSS Grade 8 association seemed to be stronger. The median-explained variance for mathematics achievement was 6.9% for mother's education, 6.7% for father's education, and 7.5% for the highest parental education (with IQRs of 6.2%, 6.1%, and 5.6%). The relationship with science achievement was comparable to that for the mathematics domain. The median-explained variance in science achievement for mother's education was 6.2%, for father's 6.9%, and for parents' highest education 7.7% (IQRs of 6.0%, 6.7%, and 6.6%, respectively). As was the case for PISA, the associations differed quite markedly across the participating countries.

6.5 Parental Occupation

(Result Tables A.16, A.17, and B.2)

Questions about the financial wellbeing of each student's family were administered in PIRLS and PISA. When making comparisons between these two studies, one first of all needs to take into account that their instruments were administered to different age groups. That said, both studies collected data from parents directly, which might put this constraint into perspective. But then again, administering a questionnaire to parents was a national option in PISA that was chosen by only 16 countries. In addition to collecting data from parents, PISA asked the students to provide information

about the occupations of their parents. In PISA, all information on occupation was transferred to the International Socio-Economic Index of Occupational Status (ISEI; see 3.3.3), thereby transferring the data to a (pseudo-)interval scale level.

6.5.1 Response Rate

The nonresponse rates for data on occupation and financial wellbeing in PIRLS were moderate to high, with medians ranging from 7.5% to 16.8% and high IQRs (from 11.3% to 25.2%). The lowest rate was for the self-reported financial wellbeing of the family (7.5%). The highest median amount of missing data came from the information on the mother's occupation and the father's occupation (both 16.8%). As we have already seen with the other variables of interest, nonresponse in the PIRLS home questionnaire seemed to result from unit nonresponse in some countries rather than from reasons connected to the variables themselves. Similar to information on parental education and language use, information on parents' occupation showed high nonresponse rates in several countries, some of which showed a similar amount of nonresponse (England and Scotland were two such examples, each with a nonresponse rate of approximately 50%).

In PISA, median missing rates were somewhat higher for the occupation data provided by parents than for the occupation data provided by students. Nevertheless, both lots of data had the same pattern. Information about the mother's job that we coded as ISEI was missing for many more students than was the case for information about the father's job (18.7% compared to 12.1% for student-provided data; 37.0% to 22.0% for parent-provided data). The combination of both parents' individual occupations that received the highest ISEI within the family provided the highest amount of valid data, with a median nonresponse rate that was still high at 20.4% for data from the parents (IQR of 20.8%) but low (4.3%) for data from the students (IQR of 3.8%).

The high amount of missing data for the information on parental occupation is, to some extent, due to the fact that people who do not have a regular occupation are not included on the ISEI and are therefore coded as missing data. Such people include apprentices, university students, unemployed people, housewives, and retirees. The nonresponse rates for the white-collar/blue-collar categorization of occupations are the same as for the ISEI variables because the bases of both are the same ISCO-88 codes. In the current study, nonresponse rates for information about mother's white-collar/blue-collar status were the same as those for the parental occupation data because they were derived from the same source of information.

Returning to the country-level results, we can see, once again, big differences across the nonresponse rates of countries. While information about mothers' occupation had the highest nonresponse rates, the combination of both the information from mothers and fathers yielded low nonresponse rates for at least five countries in PIRLS (Hong Kong SAR, Moldova, Poland, the Russian Federation, and Singapore) and for the majority (34) of the countries in PISA.

The nonresponse rate of the information from parents about household income in PISA was high, with a median rate of 22.0% (IQR of 31.7%). The higher amount of nonresponse for this item compared to the item about the family's financial wellbeing in PIRLS could be due to the different item formats. In PIRLS, the question was asked without any reference to absolute numbers (e.g., income) but rather in subjective relation to other families. In PISA, parents were directly asked about their income, or at least asked to select the applicable income-range bracket from among the categories provided.

6.5.2 Association with Achievement

In relative terms, PISA data from students about their parents' occupations showed patterns of nonresponse and association with achievement similar to the patterns for the same information provided by the parents themselves. In absolute values, the information provided by students on the highest parental occupation produced a slightly higher correlation with achievement (0.30 to 0.32 for the three subject domains of the students' data and 0.28 to 0.31 for the parents' data) and explained slightly more of the variance in achievement than did information from parents about their occupations (9.3% to 10.3% for the students' data and 8.0% to 9.5% for the parents' data). In general, associations with achievement were moderate to high for all variables on occupation and income in PISA. Median correlations varied between 0.22 and 0.31, and the median-explained variance ranged from 5.6% to 10.3% (IQRs of 4.1% to 7.0%). The variance included the items about parental occupation as well as the item on household income. The association between the white-collar/blue-collar status of parents and achievement was somewhat weaker, explaining 5.6% to 7.7% of the variance (IQRs of 3.3% to 5.6%). We suspected this pattern would become evident because these indicators reduce the variety of information to four categories only.

A look at the country results reveals some variation across countries. While several countries (Brazil, Bulgaria, Luxembourg, and Portugal) showed high associations with each of the three student-achievement areas for all of the occupation variables, others showed mainly low associations (Hong Kong SAR, Korea, Macao-China, and Qatar). Other countries had a mixture of low and moderate or moderate and high associations.

Correlation and regression analysis could not be completed with the PIRLS data because the information on occupation is provided at nominal scale level only.

6.6 Derived Scales and Indices

From the three studies of interest, only PIRLS and PISA derived scales and indices from separate variables. However, because the two studies surveyed different age populations and because we chose different sets of selected countries for the purposes of this paper, direct comparisons should not be made. We therefore analyzed the scales and indices from the two studies separately. The amount of missing data in the derived scales and indices results from the nonresponse rate for the source variables and from how the separate variables were combined into indices and scales. We

calculated Cronbach's alpha as a reliability coefficient, using all source variables for an index or a scale and following each study's description of how it derived the new variable.

6.6.1 The PIRLS Index of Home Educational Resources (HER)

(Result Tables A.18 and B.1)

The PIRLS Index of Home Educational Resources (HER) in 2006 used source variables from both the student and the parental data (see Section 5.1.1.2 or Foy & Kennedy, 2008c, for details). The index had a moderate nonresponse rate, with a median amount (7.8%) of missing data (IQR of 17.0%) resulting from the nonresponse to the source variables. Several countries (15 in total) had a (very) high nonresponse rate. Twelve other countries showed a low rate.

The median reliability (Cronbach's alpha) of the index across all countries was low, at 0.61[3] and an IQR of 0.09, thereby showing a relatively low level of dispersion across the selected countries. The item-total correlation analysis showed, on average across the countries and also within many (35) of them, a poor fit between the home-possession items (ASBGTA1-4) and the index. The reliabilities of the scale in Kuwait (0.37) and Qatar (0.38) were very low compared to those in the other countries. The only countries with a high reliability of the full scale were Bulgaria (0.78) and Iran (0.79).

The index showed a moderate median correlation with reading achievement (0.27) across the countries, with a relatively small dispersion (IQR of 0.08) crossnationally. The median amount of explained variance in students' reading achievement was moderate at 7.5% (IQR of 4.2%). Seven countries—France, Hungary, Iran, Poland, Romania, Singapore, and Sweden—showed a high association.

6.6.2 The PIRLS Index of Early Home Literacy Activities (EHLA)

(Result Tables A.19 and B.1)

The Index of Early Home Literacy Activities (EHLA) in PIRLS 2006 used only student measures as source variables related to literacy activities carried out with the child in the family before he or she began school (see Section 5.1.1.2 or Foy & Kennedy, 2008c). The EHLA showed results similar to those for the HER index: the nonresponse rate was moderate with a median amount of missing data of 6.9% (IQR of 10.4%). The reliability of the index (Cronbach's alpha) was a little higher compared to the HER index and already at a moderate level at 0.68 (IQR of 0.09). With a minimum reliability of 0.60 in Italy and a maximum reliability of 0.79 in Singapore, the index showed a somewhat similar reliability across countries. Changing the composition of the scale would not result in a substantial change in the median reliability given that all items correlated moderately with the scale.

The correlation of the EHLA index with reading achievement was quite low at only 0.17, and the IQR was 0.06, indicating that relatively few countries (11) yielded reasonable

3 We excluded England from the calculation of all median statistics for this index due to the low amount of valid data.

correlations. The median of the explained variance in reading achievement was equally low at only 2.8% (IQR of 2.0%). Only six countries (Austria, Chinese Taipei, Iran, New Zealand, Scotland, and Trinidad and Tobago) showed a moderate association with achievement. Romania was the only country to show a strong association.

6.6.3 The PISA Home Possessions Scale (HOMEPOSS)

(Result Tables A.20[4] and B.3)

The Home Possessions Scale in PISA (HOMEPOSS) had a very low level of nonresponse, with 0.3% missing data and an IQR of 0.7%. As these percentages already indicate, the level of nonresponse was low for every country. The median reliability of 0.72 (IQR of 0.12) across countries was at a moderate level, with 19 countries at the high level and nine countries at the low level. In terms of nonresponse and reliability, the HOMEPOSS seems to be promising in terms of quality. Changing the composition of the scale by removing one of the items from the scale would not result in substantial change in the median reliability. Still, item-total correlation analysis revealed several items that did not seem to fit well with the scale. Having access to the internet at home was the weakest item, with a low item-total correlation in all but five countries (Azerbaijan, Jordan, Romania, Thailand, and Tunisia).

The scale showed a moderate association with achievement in the three subjects tested in PISA. The median correlation coefficient for mathematics achievement was 0.29 (IQR of 0.11), and the median correlation with reading and science achievement was 0.28 (IQRs of 0.09 and 0.10). The explained variance in achievement was 7.6% for the reading domain (IQR of 5.1%). The amount of variance explained by the Home Possessions Scale was 8.6% (IQR of 6.7%) for mathematics achievement and 8.1% (IQR of 6.0%) for science achievement. In general, the strengths of the associations were similar for the three cognitive domains within each country. In Iceland, the HOMEPOSS did not explain any substantial variance in students' achievement (0.9% in mathematics, 0.3% in reading, and 0.6% in science). Eight other countries had only low associations for all three areas. However, 14 countries showed strong associations with the HOMEPOSS for all three domains.

6.6.4 The PISA Home Educational Resources Scale (HEDRES)

(Result Tables A.21 and B.3)

PISA's Home Educational Resources Scale (HEDRES) also showed a very low nonresponse rate, with the median amount of missing data at 0.6% (IQR of 0.8%). Only Qatar had a moderate nonresponse rate (5.5%). Nevertheless, the median reliability of the scale across the selected countries was quite low (0.53 and an IQR of 0.13). It was, in fact, the lowest of all the PISA-derived scales. Six countries (Bulgaria, Jordan, Romania, Tunisia, Turkey, and Qatar) showed a moderate reliability of the scale; for all other countries, the reliability was low. Changing the composition of the

4 The Cronbach alpha coefficients for the Family Wealth and Home Possessions Scales were computed according to the information provided in the PISA 2006 technical report. However, numbers may not match the ones given in the report (OECD, 2009, pp. 317–318).

scale would not result in substantial change in the median reliability. The item-total correlation was best for the items related to computer aids at home (computer for school work and educational software), with both items showing a moderate median correlation of 0.33 (IQRs of 0.14 and 0.18). All other items correlated to a low extent only with the total scale.

The correlations between the HEDRES and reading, mathematics, and science achievement were moderate, with 0.26, 0.26, and 0.25 for the three subject domains (IQRs of 0.08, 0.14, and 0.13). The median amount of variance explained by the scale in reading, mathematics, and science achievement was between 6.0% and 6.8% (IQRs from 4.4% to 7.0%). Again, the diversity of the results at the country level was considerable. Whereas 11 countries showed a strong association with all three domains, 14 countries had low associations. Nevertheless, given the low reliability of the measure in most of the countries, none of these results can be trusted.

6.6.5 The PISA Cultural Possessions Scale (CULTPOSS)

(Result Tables A.22 and B.3)

PISA's Cultural Possessions Scale (CULTPOSS) showed, yet again, a very low amount of missing data (1.5% and an IQR of 1.8%), and only four countries (Germany, Israel, Kyrgyzstan, and Qatar) with a moderate nonresponse rate, but also a low reliability of 0.60 and a relatively low dispersion across countries (IQR of 0.08). Only 10 countries yielded a moderate scale-reliability. Removing Item ST13Q10 (works of art) from the scale improved its median reliability (0.62), but not substantially so. Although showing the weakest item-total correlation amongst the three variables comprising the scale, the item nonetheless had a moderate correlation with the total scale.

The correlation with achievement was moderate: 0.25 for reading, 0.23 for mathematics, and 0.26 for science, with dispersion of 0.10, 0.09, and 0.08, respectively. The explained variance, at between 5.3% and 6.9% (IQRs of 4.1% to 5.1%), indicated some variation across countries. There were also differences between the subject areas. Whereas only three countries showed a high amount of explained variance with mathematics achievement, 13 countries showed a high amount with regard to reading achievement.

6.6.6 The PISA Index of Wealth Possessions (WEALTH)

(Result Tables A.23[5] and B.3)

The PISA Index of Wealth Possessions (WEALTH) also showed a low median reliability of 0.61 (IQR of 0.12). Four countries had high reliabilities, but the majority of countries (39) had only a low reliability. Removing one of the items from the scale did not improve its reliability. All items correlated moderately with it: the items on possessing cellphones and televisions showed low correlations in a couple of countries. As the value of the dispersion indicates, there was some variation across countries with respect to the reliability of the scale.

5 The Cronbach alpha coefficients for the Family Wealth and Home Possessions Scales were computed according to the information provided in the PISA 2006 technical report. However, numbers may not match those given in the report (OECD, 2009, pp. 317–318).

The median amount of missing data was not only low at 0.3% (IQR of 0.7%) but also low for every country. The association with achievement was, however, weak: for reading it was 0.12, mathematics 0.17, and science 0.14. Moreover, the IQRs for these correlation coefficients showed quite high dispersion across the countries: the values of the IQRs were around the same value as the correlation coefficients (0.13 for all three domains). The amount of explained variance for reading was 1.6%, for mathematics 2.8%, and for science 2.1%, and the corresponding dispersions, when compared to the absolute values of the medians, were somewhat high at 4.6%, 2.8%, and 3.9%. Only two countries (Brazil and Chile) had a high amount of explained variance for all three subject domains. However, 12 countries had almost no explained variance (less than 1.0%) for any of the three subject areas.

6.6.7 The PISA Index of Economic, Social, and Cultural Status (ESCS)

(Result Tables A.24 and B.3)

The reliability for the Index of Economic, Social, and Cultural Status (ESCS) was moderate (0.68 and an IQR of 0.10). Because the index is composed of three scales that represent three different constructs, we could expect the reliability to increase if we removed one of the scales from the index. The scales comprising the index correlated moderately (HISEI and HOMEPOS) and even highly (PARED) with the index. There was some variation across the countries with respect to the reliability of the ESCS index: 18 countries showed a low reliability and five countries a high reliability.

The amount of missing data was low, with a median nonresponse rate of 0.6% (IQR of 0.8%). Only Israel showed a moderate rate of nonresponse (5.2%). The correlations of this index with achievement were not only high for all three subjects but also the strongest of all the indices and scales for both studies, with a median correlation of 0.35 for reading, 0.37 for mathematics, and 0.37 for science, and IQRs of 0.09, 0.08, and 0.10, respectively. This pattern also applied to the amount of the variance explained by this variable, with median variances of 12.0% for reading achievement (IQR of 5.7%), 13.9% for science (IQR of 7.2%), and 13.6% for mathematics achievement (IQR of 6.2%). This finding was not unexpected because the index combines information and explanatory power from three sources derived from many separate variables (the index of home possessions, the highest ISEI, and parental education, defined in terms of years of schooling). Most countries showed a strong association with each of the three areas of achievement at (well) above 10.0% of the explained variance. The only countries where the associations were low for all three areas were Macao-China and Qatar.

7. Discussion

Before discussing the results of this study in detail, we consider it worthwhile to begin by pointing out some general tendencies. The overall impression is that data obtained from parents in both PIRLS and PISA tended to have a higher amount of nonresponse compared to the data provided by students. In some cases, the median nonresponse rate in the parental data was high, at about 15%. In one case (occupation of mother in PISA), the amount of missing data was about 40%. This exceptionally high rate had two origins. First, the nonresponse rate was again about the same size as that for other information collected from parents (less than 20%). Second, about 20% of the parents answered that the mother was working as a housewife and therefore did not have a paid occupation. In this case, missing data could be compensated for by using the information pertaining to both the mother's and the father's occupation as a combined indicator for the family (e. g., the highest ISEI score derived from the occupations coded using the ISCO 88). This way, only 2% of the students were left with missing data for the combined data in addition to the unit nonresponse.

The reasons why information collected from parents consists of a much higher nonresponse rate could be many and different, ranging from lack of interest to unwillingness to complete the questionnaire or separate parts of it. Some parents may not have been willing to provide information they considered too sensitive (private). Given the higher nonresponse rates of parents compared to students, one could argue about whether there really are benefits to including parents as a source of information in large-scale assessment studies. First of all, the extra costs associated with developing, administering, and processing the data from an additional parental questionnaire are considerable. Secondly, research shows that information from Grade 8 and/or 15-year-old students about their families tends to be as valid and reliable as the information provided by their parents (see Chapter 3.5). Nevertheless, because parental occupation has been identified as an important indicator of family background, collecting this data seems a worthwhile exercise. For the older age group in the studies of interest (Grade 8 and/or 15-year-old students), information on parents' occupation could be collected from the students, whereas for the younger age group (Grade 4 students) the preferred option appears to be that of obtaining the information directly from parents. The same recommendation seems applicable to information on parents' education.

The different domains of family background (home possessions, parental education, etc.) showed different strengths of associations with achievement across subject domains and studies. Of all the indicators that we analyzed, the frequency with which the language of the test was spoken at home (the only nondichotomous language variable) showed the weakest relationship with achievement for all subjects amongst all domains.

7.1 Home Background Domains

7.1.1 Immigration Status

The variables related to immigration status generally showed a weak association with achievement. However, in TIMSS Grade 4, the median amount of variance explained by whether or not the students were born in the country of the test was moderate for both mathematics and science. Parents' immigration status seemed to play only a minor role. In regard to the younger group (the Grade 4 students), the information on whether the student was born in the country of the test had more explanatory power than the information on whether his or her parents were born in the country of the test. This pattern was similar for the TIMSS Grade 8 data, but the absolute strength of both associations was much lower than for those emerging from the TIMSS Grade 4 data.

The PISA data showed barely any relationship between the immigration status of students and achievement. For the majority of the countries participating in PISA 2006, the association with achievement was low, although in some countries the regression results were statistically significant, suggesting that immigration status might play a role with respect to the achievement of the younger age group (at least for TIMSS Grade 4 students) but not of the older age group (PISA and TIMSS Grade 8 students). In PISA, the median relationship between the age at which a student immigrated to the country of the test and achievement was negative. Thus, the older students are when they immigrate, the lower their achievement tends to be.

This finding could be simply explained by the fact that students might need to learn the language first before being able to follow instructions in school, with the latter becoming more complex and demanding with higher grade level because of the increasing difficulty of the content taught. However, we cannot regard this finding as a general tendency because, in the majority of countries, the correlation between the variable and achievement in the different subjects varied considerably across the countries and was very weak (below 0.10 in absolute terms). In some countries, it was even positive, but again very weak.

Nevertheless, immigration status is, and likely will remain, a prominent issue on the political agenda, mainly because it is one of the major aspects of equal opportunities in education. International comparative research in education should continue to include the immigration status of students as a characteristic of family background. Even more, this characteristic needs to be more closely integrated into theoretical frameworks as, for example, aspects of social and cultural capital. This approach, in turn, means giving consideration to the role immigration policies in countries might

play in trying to counteract the types of problems students can face when settling in a new country. The possibilities of using immigration status as a component of a scale or an index are also worthy of investigation.

7.1.2 Occupation

In PISA, all occupational variables showed moderate to strong associations with achievement for all three subject domains. The variables on parental occupation as measures of family background seemed to have the largest impact on the educational outcomes. As Hauser (1994) argued, parental occupation is one of the core components of SES (and also of family background) and, compared to other measures, is much more stable over time. As such, it can be used as a long-term indicator of family income. Although the association between the occupational variables and achievement was substantial, the issue of missing data still needs to be considered. In PIRLS, the parent-reported occupational data had an equal amount of missing data for both parents. In PISA, the amount of missing data was at least twice as high for the mother's compared to the father's job, regardless of whether students or their parents provided the information.

The reason seems to originate mainly from question format. The multiple-choice format used in PIRLS has the advantage of respondents only needing to tick an answer. The effort to write in a job title and a job description, as is the case in the PISA questionnaires, is considerably more onerous, which could be one reason for not answering. Also, providing details about an occupation, as in PISA, requires more knowledge compared to only broadly classifying an occupation, as in PIRLS. On the one hand, respondents might be ashamed about their lack of exact knowledge and decide not to provide any information, even though providing just some broad information could still be useful in terms of classifying the occupation. On the other hand, although PIRLS had lower amounts of missing data on occupational variables, the PISA data were much more detailed, and the ISCO classification scheme could be applied with several hundred categories, providing a much finer classification compared to the nominal 10-point scale used in PIRLS. Finally, the ISCO codes can be recoded into the metric scale of the ISEI. In sum, we consider that the several advantages of the open-ended question format for information on occupation outweigh the advantages of a closed question format.

7.1.3 Parental Education

The association of parental education with achievement for the older age group (TIMSS Grade 8 and PISA) was moderate for all tested subjects. The associations were a little higher for TIMSS, but the differences were very small. It can be argued that of all the domains of family background discussed so far, the variables on parental education together with occupation were the strongest predictors of achievement. Since the introduction of the ISCED classification (see Chapter 3.2.10), the administration of parental education in international comparative education research has become a standard procedure. Several studies have published examples showing how adaptation of the abstract ISCED education levels to qualifications specific to national education systems facilitates international comparisons (Foy & Kennedy, 2008b; Foy & Olson,

2009b; OECD, 2009). Parents' education can also serve for studying the reproduction effect proposed by critical theorists; parents' education consequently has also been incorporated into combined measures of family background. Parental education should thus remain a standard characteristic of family background in international large-scale assessments. Still, the international classification scheme used for parental education, that is, the ISCED classification, should undergo constant examination concerning its best possible fit to countries' education systems.

7.1.4 Home Possessions

We operationalized the domain of home possessions in the three large-scale studies of interest by using the highest number of items amongst all five domains of family background. However, only a few of the items showed substantive association with achievement. In all three studies, number of books at home appeared to be the strongest predictor of achievement. It certainly had a strong association with achievement across the different studies and subject areas investigated. This pattern also applied to the association between number of children's books at home and reading achievement in PIRLS.

In general, the existence of educational aids at home showed a low or even almost no association with achievement. The same applied to items of everyday life, such as dishwashers, cars, televisions, and cellphones. The availability of a computer at home showed a moderate amount of explained variance in PISA for all subject areas, but had less explanatory power with respect to the TIMSS Grade 8 data. In both PIRLS and TIMSS Grade 4, the amount of explained variance in achievement on any subject was small. The association between the family having an internet connection and achievement was also low.

These results suggest that computers play a different role for the different age groups. Students from the younger age group might have been using computers at home much less frequently or for different purposes than were the students from the older age group. Use of computers (e.g., as educational aids) thus becomes an additional aspect of computers at home. This explanation would leave the possession of a computer at home as having lesser importance as an indicator of family wealth.

In general, all country-specific (optional) home possession items showed only a very small amount of explained variance in achievement. Their usefulness in international comparisons is questionable not only because of the low association with achievement, but also because each of the countries participating in the IEA and OECD studies was allowed to choose its own set of possession items, thus compromising comparability. In addition, each one of the optional items had its own value for the specific culture, but the extent to which items with different values compare across countries is not clear from the study reports.

7.1.5 Scales

The reliability of the derived scales (PIRLS and PISA) was moderate for four of the scales and rather low for the remaining three. In general, a number of variables showed a strong association with achievement across the three studies and different

age groups, but not all background domains and single items within the domains showed as strong a relationship as commonly described in the literature. The indices in PIRLS showed satisfactory quality overall, although the reliability of the Index of Home Educational Resources was relatively low (less than 0.65). Nevertheless, this variable had a moderate association with reading achievement. In contrast, the Index of Early Home Literacy Activities had moderate reliability, but the association with achievement was quite weak. For the Index of Home Educational Resources in PISA, the median reliability was quite low, although it showed a strong association with achievement. The Index of Family Wealth Possession in PISA had an almost moderate median reliability, but the association between this variable and achievement was the weakest of all the PISA indices. The highest reliabilities and associations with achievement in PISA that we found were with the Home Possessions Scale and the Index of Economic, Social, and Cultural Status.

7.2 Recommendations

Based on the analyses that we conducted, we offer several recommendations regarding the measurement of family background:

1. With respect to all home-possession variables, we recommend including those that show the highest association with achievement in terms of explained variance, namely, number of books in the home, number of children's books in the home, number of student's own books, and access to a computer. This list applies to all three studies—TIMSS, PIRLS, and PISA. Regardless of a strong desire by single countries to include certain home-possession items (and there might be good reasons to include these on a national level), we suggest some standard for a minimum level of discrimination. Usually, items regarded as important will enter the field trial phase, with efforts then made to verify their appropriateness for inclusion in the main survey phase, the outcomes of which are, of course, reported at a later date. A standard for minimum discrimination of families' home-possession items could enhance opportunities to include items that bear the potential to explain variance in student achievement outcomes.
2. Although TIMSS does not collect occupational data, occupational variables have shown their importance and explanatory power in other studies and also in ours. Single items on employment situation in PIRLS showed only a weak association with achievement. Still, constructing a scale together with other occupation measures might yield a better explanation of results. Because collecting information about parental occupation from Grade 4 students does not seem to be reasonable, future TIMSS cycles might consider using a home questionnaire for that grade. This, of course, would introduce extra costs for both international coordination as well as participating countries in terms of developing, administering, and processing an additional questionnaire. A concern might still be the high rate of nonresponse seen in PISA and PIRLS. However, as the results from the analysis reported in this paper have shown, detailed information about occupation can be fruitfully included in models for explaining education outcomes. In TIMSS Grade 8, the information about parental occupation could be collected from students,

given that other studies have already shown that students at this age can provide reliable data about their parents. Thus, information with reasonable explanatory power could be collected without additional questionnaire development-and-administration costs. Only some additional processing would be necessary, mainly in terms of coding occupation information into classification schemes such as the ISCO.

3. Although PIRLS collects parental occupation data, it does so only on a nominal scale level. PIRLS could therefore also profit from choosing to collect more detailed information on parents' occupation. A finer indicator would be desirable so that the effect of family background on achievement could be analyzed in more depth. Using an open format and coding the answers using the ISCO classification scheme seems to us to be a reasonable option that could result in a metric occupation scale, such as the ISEI. An example from Caro (forthcoming) shows that this lack of detailed information is a drawback of PIRLS because it limits the classification of the collected occupational data in other scale types, such as ISEI, even for research purposes.
4. Should the three studies decide to collect more detailed occupation data, we recommend that they include questions about parents' self-employment status and the number of people they supervise in their work. This would allow derivation of EGP classes and, from there, investigation of their association with the assessment outcome variables. The usefulness of these classes was just one example shown by Baumert and Schümer (2001) when they reported the relationship between EGP class and reading achievement in Germany, using data from the PISA 2000 study.
5. PISA derives several scales of aspects of family background, but the reliability tends to be relatively low for the scales on home educational resources, on cultural possessions, and on family wealth. A review of the composition of the scales and further research on optimized operationalizations of aspects of family background might facilitate improvements to these measures of family background. The family wealth scale, in particular, seems to function somewhat differently in different countries. At first glance, the scale seems to relate to the country's economic development, with Western industrialized countries such as Australia, Norway, Slovenia, and Sweden showing no association between family wealth and student achievement, and less-developed countries such as Colombia, Thailand, and Tunisia showing the highest association among the countries in scope. But then the United States also shows up among the countries with the highest association, and Chinese Taipei among the countries with the lowest association. The country-related aspects of family wealth indicators thus need to be investigated further.
6. TIMSS does not derive any scales or indices of family background, although in general these derived variables have shown their predictive power regarding students' achievement. The creation of such indices in future TIMSS cycles might be considered, given they have shown to be highly beneficial for research purposes. Such scales have already been created by researchers working with data from the large-scale assessment studies. For example, May (2002) conducted research which showed that it is possible to create a reliable, valid, and internationally

comparable SES scale for TIMSS and to use it successfully in analyses of study data. Similar findings come from Caro (forthcoming), who constructed a SES scale for PIRLS, using the available relevant data from the study. The derived SES scale had a satisfactory correlation with achievement and satisfactory reliability across countries, although the crossnational validity analysis yielded unsatisfactory results.

Deriving scales from single items offers several advantages. First, reporting can be linked more directly to latent constructs described as important in a study's framework. Second, scales or indices comprised from several items bear the potential of producing a better representation of characteristics of interest with regard to cultural differences among different societies and countries around the world. For example, the cultural value of items is likely to differ a lot between societies when cultural capital is operationalized (maybe amongst other indicators) as home possessions, thereby leaving the value of reporting on single items in doubt. But a combined scale could account for differential value ascriptions of different home possession items in different cultural environments. Of even more pertinence, a combined scale could comprise different (sets of) home-possession items for different cultures. Third, a simple practical reason favors reporting on scales rather than on item level. When deriving valid data for a respondent's scale score, it is usually not necessary to have all items contributing to the scale with valid data. Thus, even when there are (some) missing data in the single items' data, it could be possible to assign valid scale values/scores to more respondents than if any of the single items or a simple combination of the items were used. The latter would necessitate the application of listwise deletion in the analysis.

A further recommendation can be derived from other research that has used recent data from the TIMSS and PIRLS cycles. Both the TIMSS and PIRLS research teams might consider constructing background indices using more complex methods, following examples from the likes of Van Damme, Liu, Vanhee, and Pustjens (2010), who derived a SES index for PIRLS that showed good quality and accounted for a high amount of explained variance in achievement. In her recently published dissertation, Preuschoff (2010) outlined the potential of using Rasch scales for analyzing and reporting TIMSS and the PIRLS background data, in particular effective learning environments connected to the TIMSS 2011 framework (Mullis, Martin, Kennedy, Trong, & Sainsbury, 2009) and PIRLS 2011 framework (Mullis, Martin, Ruddock et al., 2009). She showed that complex scales can be derived with good-quality statistics from the already existing items used in the two studies. In general with Rasch scaling, the respondents can be put on the same metric as the items. Accordingly, "if the content of an item can be paired with its location on the map for the scale, the Rasch scale can be conveniently described in a way that gives meaning to a country's position on the scale" (Preuschoff, 2010, p. 8). Another advantage of analyzing scales instead of single items is attainment of a higher level of measurement. In most cases, an interval scale level can be achieved with scales or indices instead of the ordinal or even nominal scale level of most of the single items. In short, more analytical techniques and computational operations are available to researchers than might appear to be the case.

As we mentioned in Chapter 3.1.4, the SACMEQ approach to comparing findings could be used to improve the understanding of country differences with regard to students' achievement. By comparing absolute values of socioeconomic background across countries (as opposed to relative measures such as percentiles within a country), the relationship between a country's distribution of student achievement and family background differences within the country can be compared with the same relationships in other countries. Future cycles of TIMSS, PIRLS, and PISA could adopt this approach to provide an alternative view of differences in student achievement across countries.

The research by Gershoff, Aber, Raver, and Lennon (2007; see also Chapter 3.2.2 of this paper) suggests a mediating effect of family hardship on the association of income with cognitive skills. If the issues with measuring income can be dealt with to the extent that good-quality data are collected in terms of response rates and validity, it might be worth checking for this type of mediating effect in future large-scale student assessments. However, there might still be issues with measuring family hardships at the international level as, for example, with the definition of poverty, which is a constantly discussed topic on the political agenda and which is frequently changed within countries to suit political purposes.

As we pointed out in Chapter 3.2.8, TIMSS and PIRLS include school-level measures pertaining to the socioeconomic composition of the student body. Information about the proportion of students from economically disadvantaged homes does not provide information about the neighborhood of the student's home, per se, but researchers could think about using it as a proxy for the neighborhood composition of the student's home, given there is only one school that all young people from a certain area attend (because of the lack of alternative schools in the vicinity). Still, this assumption might only hold for rural areas, and, even then, parents in these areas might be able and willing to spend time and money to send their children to a different school further away. It therefore seems desirable to collect information about the neighborhood of the students' home directly by, for example, asking the parents. Of course, the validity of this information needs to be checked, as there might be an issue with respondents giving answers that they see as socially desirable. Some parents, for example, might feel ashamed to reveal that their family is living in a neighborhood with a bad reputation.

We also suggest paying closer attention to religion as a measure of students' family backgrounds. As we mentioned in Chapter 3.2.9, IEA's ICCS survey includes engagement with religion as a part of broader civic engagement in order to investigate the attitudes of students toward the influence of religion in society. It would be interesting to explore if religion is associated not only with students' attitudes but also with students' achievement.

A final suggestion is to further explore practices within families. PIRLS and PISA both include a few questions about activities in the family directed toward improving the child's reading (PIRLS) and science (PISA) skills. Research, we think, should focus more on family practice with respect to supporting students' skills development.

7.3 Further Research Needs

It was not possible for us, in the current project, to cover all issues related to the measurement of home background in large-scale international assessment studies of educational achievement. The main purpose of our study was to explore some basic issues related to family background measures—nonresponse rate, degree of association with student achievement, and reliability of the scales and indices. And even though these basic characteristics address only some of the issues concerning the measurement of the family background of students, they are still ones that merit ongoing consideration.

In this paper, we briefly addressed the possible role of countries' immigration policies with regard to the effect of the immigration status of the family on students' achievement. Further research could take a closer look at the similarities and differences between the immigration policies of the countries participating in international large-scale assessments.

It would also be desirable to investigate further the use of indicators of family background on a higher aggregated level such as neighborhood characteristics. For example, TIMSS 2007 reported on information provided by school principals about the percentages of students in their respective schools who were from economically disadvantaged families and/or had the language of the test as their native language (Mullis et al., 2008). The inclusion of such variables in a more complex model (multilevel) could reveal the contribution that these variables make in terms of explaining differences in student achievement. Multilevel analyses (e. g., multilevel regression analyses or hierarchical linear modeling) could be used to further explore critical factors of the family's environment, on the one hand, and to differentiate the influence of the factors from different levels (e. g., student, school, and country), on the other. Examples of country-level data that are available from diverse sources could include educational expenditure, achievement orientation in general, and the specific subject domain that is the focus of the respective student assessment.

Another direction relative to more advanced methods of exploring the influence of family background on student achievement was recently proposed by Sandoval-Hernández (2012). The author used exploratory structural equation modeling (ESEM) to explore the factor structure of indicators of economic, social, and cultural capital as operationalized in PIRLS 2006 and PISA 2009. ESEM allows for items to load on more than one factor simultaneously. Sandoval-Hernández's results suggest that more sophisticated analyses of this kind (and which were beyond the scope of our paper) could yield more insight into the structure of family background items.

Finally, we draw attention to two additional important issues related to family background that we consider are very much in need of further research-based exploration. They are the crosscultural validity of the family background measures and their coverage in the studies. We therefore end this paper by sketching further research needs with regard to those two aspects.

7.3.1 Crosscultural Validity

Addressing the measurement of family background in international large-scale assessment studies requires analysis of crosscultural validity that is structure-oriented or, in other words, concerned with determining whether the measures (in our case) of family background used in TIMSS 2007, PIRLS 2006, and PISA 2006, or of any other large-scale assessment study measure the same construct across all countries (see Van de Vijver, 2003a).

We mentioned in Chapter 3 that comparing home possessions across countries is not a straightforward process. Home possessions are not the only domain in which crossnational comparisons can be problematic. According to Braun and Mohler (2003), national characteristics, typical social structures, and legal institutions can make the definition of a latent construct quite a difficult task, and the indicators chosen may not be equivalent across countries and cultures. The same is true for the questions in survey instruments. All of this can result in a comparison of incomparable things (Braun & Mohler, 2003).

The measurement of a construct across countries is usually described with either one of two terms—"cross-cultural bias" and "cross-cultural equivalence." These two terms are generally regarded as antonyms. On the one hand, bias refers to the "presence of nuisance factors that challenge the comparability of scores across cultural groups" (Van de Vijver, 2003a, p. 144). The notion of biased scores lays the focus on cultural differences. On the other hand, equivalence refers to the comparability of scores across cultures. Bias and equivalence are not related to the survey instrument itself, but rather to its applications in different countries. For this reason, the presence or absence of bias or equivalence needs to be determined empirically.

Sources of bias can also reside in the constructs themselves, the methods used, and the items by which the constructs are measured; hence, three different types of bias can be distinguished (see Van de Vijver, 2003a, pp. 145–147). They are construct, method, and item bias.

Construct bias is inherent when the measured construct is not identical across cultures. Sources of construct bias could include the following:

- (Partial) differences in the meaning of a construct definition across cultures;
- Different behaviors associated with the construct across cultures;
- Poor selection of the behaviors that are manifesting the construct (i.e., dimensions of the construct measured by the items in the instrument); and
- Incomplete coverage of the relevant aspects of the construct itself.

Sources of method bias typically include (amongst others):

- Samples that cannot be compared with each or one another;
- Ambiguous instructions for the respondents or survey administrators; and
- Differences in familiarity with the stimulus material or response procedures.

Item bias can occur because of:

- Ambiguous items or poor translations;
- Nuisance factors (e. g., an item invoking additional traits or abilities); and
- Culture-based peculiarities (Van de Vijver, 2003a).

However, scores and scales should not depend on variance that is not relevant to the underlying construct (French & Finch, 2006). As Schulz (2006) points out, measurement of family background of students across countries requires crosscountry validation of the underlying construct. A high nonresponse rate also introduces bias. Keeves, Lietz, Gregory, and Darmawan (2006) point out that, in the case of achievement, for example, bias due to nonresponse inflates the mean level of performance and reduces the variance. This kind of bias would give erroneous estimates of the achievement and also reduce the capability of the analysis of variance techniques used in later analyses (Keeves et al., 2006). As Mullis (2002) also points out, the nonparticipation of students introduces bias in the results by increasing or decreasing the performance. Nonresponse and nonparticipation can also be driven by cultural differences and thus need to be investigated with regard to crosscultural validity.

The aforementioned issues pose a problem of marked current relevance for international comparative studies. Measures and indicators need to be comparable across participating countries and/or education systems. When reporting relations or correlations at the international level, researchers need to be sure that indicators are measuring the same things across countries. The validity of scales (indices) obtained by international large-scale assessment studies (such as TIMSS 2007, PIRLS 2006, and PISA 2006) need to be analyzed. Such an analysis would include only those scales that show a promising quality in terms of nonresponse and reliability, as explored in the previous stages of this paper.

Bias or equivalence analysis endeavors to determine the level of comparability of data across cultures. It also verifies whether nuisance factors are distorting the results. The analysis of bias in crosscultural studies attempts to identify whether nuisance factors associated with variation are present. Equivalence analysis is aimed at determining the consequences of bias on crosscultural comparisons.

Different statistical techniques can be used to explore whether the underlying construct measured is the same across cultures. Researchers wanting to determine crosscultural validity could use multigroup confirmatory factor analysis (MGCFA). Confirmatory factor analysis (CFA) is "one of the several statistical techniques that form part of structural equations modeling" (Van de Vijver, 2003b, p. 212). CFA involves decomposing correlations or covariance and then testing to what extent the observed covariance can be reconstructed. This approach is carried out while assuming a specified (in advance) factor constellation. The specific parameters that need to be estimated can be constrained to be equal, after which the factor loadings, factor covariance, and error variances can be examined (Van de Vijver, 2003b). Constraints

for reviewing the invariance of model parameters can be of different types once the invariance of the factor structure and factor loadings have been examined (Schulz, 2006).

Other methods and models that are often applied in crosscultural validity studies use item response theory (IRT) as a basis. IRT has some attractive features, such as the link between the person and item parameters, which produces results that can be empirically examined. Also, once data have been fitted, IRT allows comparisons of item parameters across cultures. This can be done by using statistical tests of differences on these parameters, and those tests can often then be used to identify items that are biased. In addition, IRT can deal with instruments that are not identical in all countries. Item parameters, however, can be compared across countries regardless of the differences in the scores in each one of them, but only if the underlying latent trait (construct) in all countries is the same (Van de Vijver, 2003b).

This brief overview of the two statistical methods shows that a crosscultural validity analysis of all derived scales and indices used as a measure for student family background would be desirable. It also shows multigroup confirmatory factor analysis or multigroup IRT to be methods that are well suited to this process.

7.3.2 Coverage

The concept of family background refers to many aspects that cannot be covered entirely within one study. There are too many questions that theory states to be important, on the one hand, and too little time and too few resources to enable one to attempt to answer all those questions, on the other. Nonetheless, an important part of future research would be to identify the aspects of family background encompassed by the international large-scale education studies for predicting the achievement of the students assessed. Moreover, this analysis should include not only a listing of the aspects that were actually assessed, but also a side-by-side comparison of the studies with regard to the aspects covered and the number and types of variables used. Content analysis of the assessment frameworks and of the instruments that were used to collect contextual data should be the methodological approach for this kind of research. Such an analysis would help to delineate the differences in the approaches that the large-scale studies that we have considered in this paper have used to measure students' family backgrounds.

Content analysis is a method that has long been used in research. According to Krippendorff (2004, p. 18), content analysis is "a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use." Although some authors argue that content analysis is a quantitative method, Krippendorff (2004) distinguishes qualitative and quantitative alternatives of the technique, and gives preference to the qualitative ones—"Reading is a fundamentally qualitative process, even when it results in numerical results" (Krippendorff, 2004, pp. 19–20)—and states that replacing the exact quoting with numerals is just for convenience (Krippendorff, 2004, p. 87). He then advises identifying and counting the frequency of occurrence of each of them.

This is the reason why the boundary between qualitative and quantitative content analyses is blurred (Priest, 2009, p. 40). Quantitative approaches reduce the complexity of information. This process involves isolating single elements of the information, counting up their specific characteristics, and then classifying them against criteria. Qualitative approaches attempt to identify the meaning of the information as a whole. Identifying the meaning by interpretation is the next step (Kelle, Prein, & Bird, 1995, p. 168). While a distinctive trait of qualitative and quantitative content analyses is their orientation to manifest (in quantitative) and latent (in qualitative) content, either kind of approach often means analyzing the traits ascribed to the other. To put this consideration another way, sometimes the qualitative approach deals with the manifest content and the quantitative with the latent. Nevertheless, the qualitative content analysis is concerned more with the latent content and "can [therefore] better take into account subtleties of the structure of arguments and narratives not easily captured by quantitative summaries" (Priest, 2009, p. 108).

Future research intent on measuring family background from a methodological perspective and focusing in particular on the approaches adopted by international large-scale education studies should use a mixture of qualitative and quantitative variants of content analysis, but pay greater heed to the qualitative. The indicators of family background and the variables derived from them (indices, scales) could then be classified according to the aspects of family background they cover in each of the studies. The studies could then be compared and evaluated with regard to the presence or lack of variables and indicators in each of the aspects that comprise the theories of family (home) background.

References

Adams, R., & Wu, M. (Eds.). (2002). *PISA 2000 technical report*. Paris, France: Organisation for Economic Co-operation and Development (OECD).

Alison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.

Baratz-Snowden, J., Pollack, J., & Rock, D. (1988). *Quality of responses of selected items on the NAEP special study student survey*. Princeton, NJ: National Assessment of Educational Progress.

Baumert, J., & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenz- erwerb [Family living conditions, participation in education, and skills acquisition]. In Deutsches PISA-Konsortium (Eds.), *PISA 2000: Basiskompetenzen von Schüler-innen und Schülern im internationalen Vergleich* [PISA 2000: Basic skills of students in national and international comparison] (pp. 323–407). Opladen, Germany: Leske & Budrich.

Bond, G. C. (1981). Social economic status and educational achievement: A review article. *Anthropology & Education Quarterly*, *12*(4), 227–257.

Bornstein, M. H., & Bradley, R. H. (Eds.). (2003). *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bourdieu, P. (1977). Cultural reproduction and social reproduction. In J. Karabel & A. H. Halsey (Eds.), *Power and ideology in education* (pp. 487–511). New York, NY: Oxford University Press.

Bourdieu, P. (1986). The forms of capital. In J. E. Richardson (Ed.), *Handbook of theory of research for the sociology of education* (pp. 241–258). New York, NY: Greenwood Press.

Bowles, S. (1977). Unequal education and the reproduction of the social division of labour. In J. Karabel & A. H. Halsey (Eds.), *Power and ideology in education* (pp. 137–153). New York, NY: Oxford University Press.

Brasington, D., & Haurin, D. R. (2006). Educational outcomes and house values: A test of the value added approach. *Journal of Regional Science*, *46*(2), 245–268.

Braun, M., & Mohler, P. (2003). Background variables. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 101–116). Hoboken, NJ: John Wiley & Sons.

Buchmann, C. (2002). Measuring family background in international studies of education: Conceptual issues and methodological challenges. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 150–197). Washington DC: National Academy Press.

Campbell, J. R., Kelly, D. L., Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001* (2nd ed.). Chestnut Hill, MA: Boston College.

Caro, D. (forthcoming). *Measuring family socioeconomic status: A methodological proposal for the Progress in International Reading Literacy Study*. Paper prepared for the fourth Conference of the European Survey Research Association (ESRA), Lausanne, Switzerland.

Coleman, J. S. (1985). Achievement level, social class, and the self-concepts of mildly handicapped children. *Journal of Learning Disabilities*, *18*(1), 26–30.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*(Supplement), 95–120.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L., (1966). *Equality of educational opportunity*. Washington, DC: National Center for Educational Statistics.

Cortina, J. M. (1993). What is coefficient alpha? Examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104.

Cronbach, L. J. (1960). *Essentials of psychological testin*g (2nd ed.). New York, NY: Harper & Brothers.

De Graaf, N. D., De Graaf, P. M., & Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: A refinement of the cultural capital perspective. *Sociology of Education*, *71*(2): 92–111.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Dika, S. L., & Singh, K. (2002). Applications of social capital in educational literature: A critical synthesis. *Review of Educational Research*, *72*(1), 31–60.

Duncan, G. J., & Magnuson, K. A. (2003). Off with Hollingshead: Socioeconomic resources, parenting, and child development. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 83–106). Mahwah, NJ: Lawrence Erlbaum.

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, *15*(1), 35–54.

Elley, W. B. (1992). *How in the world do students read?* IEA Study of Reading Literacy. The Hague, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Ensminger, M. E., Forrest, C. B., Riley, A. W., Kang, M., Green, B. F., Starfield, B., & Ryan, S. A. (2000). The validity of measures of socioeconomic status of adolescents. *Journal of Adolescent Research*, *15*, 392–419.

Entwisle, D. R., & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development*, *65*(6), 1521–1540.

Erikson, R., & Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford, UK: Clarendon Press.

Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, *30*, 415–451.

Everitt, B. S. (2002). *The Cambridge dictionary of statistics* (2nd ed.). Cambridge, UK: Cambridge University Press.

Everson, H. T., & Millsap, R. E. (2004). Beyond individual differences: Exploring school effects on SAT scores. *Educational Psychologist*, *39*(3), 157–172.

Fetters, W., Stowe, P. S., & Owings, J. A. (1984). *High school and beyond: A national longitudinal study for the 1980s. Quality of responses of high school students to questionnaire items* (NCES 84-216). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Foy, P., & Kennedy, A. M. (Eds.). (2008a). *PIRLS 2006 user guide for the international database*. Chestnut Hill, MA: Boston College.

Foy, P., & Kennedy, A. M. (Eds.). (2008b). *PIRLS 2006 user guide for the international database: Supplement 2. National adaptations of international background questionnaires and achievement items*. Chestnut Hill, MA: Boston College.

Foy, P., & Kennedy, A. M. (Eds.). (2008c). *PIRLS 2006 user guide for the international database: Supplement 3. Derived background variables*. Chestnut Hill, MA: Boston College.

Foy, P., & Olson, J. F. (Eds.). (2009a). *TIMSS 2007 user guide for the international database*. Chestnut Hill, MA: Boston College

Foy, P., & Olson, J. F. (Eds.). (2009b). *TIMSS 2007 user guide for the international database: Supplement 2. National adaptations of international background questionnaires*. Chestnut Hill, MA: Boston College.

Freidlin, B., & Salvucci, S. (1995). *Empirical evaluation of social, psychological, & educational construct variables used in NCES surveys* (Working Paper No. 95-14). Washington, DC: U. S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling*, *13*(3), 378–402.

Gamoran, A., & Long, D.A. (2006). Equality of educational opportunity: A 40 year perspective. In R. Teese, S. Lamb, & M. Duru-Bellat (Eds.), *Education and equity: International perspectives on theory and policy* (pp. 3–27), New York, NY: Springer.

Ganzeboom, H. B. G., De Graaf, P., & Treiman, D. J. (1992). A standard International Socio-Economic Index of Occupations. *Social Science Research*, *21*(1), 1–56.

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, *25*(5), 201–239.

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development*, *78*(1), 70–95.

Gorman, S., & Yu, C. C. (1990). *Science achievement and home environment: National Assessment of Educational Progress 1985–1986*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Gottfried, A. (1985). Measures of socioeconomic status in child development research: Data and recommendations. *Merrill-Palmer Quarterly*, *31*(1), 85–92.

Hatcher, R. (1998). Class differentiation in education: Rational choices? *British Journal of Sociology of Education*, *19*(1), 5–24.

Hauser, R. M. (1994). Measuring socioeconomic status in studies of child development. *Child Development*, *65*(6), 1541–1545.

Hauser, R. M., & Warren, J. R. (1997). Socioeconomic indexes for occupations: A review, update, and critique. *Sociological Methodology*, *27*, 177–298.

Hobsbawm, E. J. (1972). Class consciousness in history. In I. Meszaros (Ed.), *Aspects of history and class consciousness* (pp. 5–21). London, UK: Routledge and Kegan Paul.

Hoxby, C. M. (2001). If families matter most, where do schools come in? In T. Moe (Ed.), *A primer on American schools* (pp. 89–125). Stanford, CA: Hoover Institution Press.

International Labour Office (ILO). (1968). *International Standard Classification of Occupations: ISCO-68* (rev. ed.). Geneva, Switzerland: Author.

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

Joseph, S. (1994). Gender and family in the Arab world. In S. Sabbagh (Ed.), *Arab women: Between defiance and restraint* (pp. 194–201). New York, NY: Olive Branch Press.

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, *29*(1), 417–442.

Keeves, J. P., Lietz, P., Gregory, K., & Darmawan, I. G. N. (2006). Some problems in the analysis of cross-national survey data. *International Educational Journal*, *7*(2), 110–126.

Kelle, U., Prein, G., & Bird, K. (1995). *Computer-aided qualitative data analysis*. Thousand Oaks, CA: Sage Publications.

Kohen, D. E., Brooks-Gunn, J., Leventhal, T., & Hertzman, C. (2002). Neighborhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development*, *73*(6), 1844–1860.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Kurki, A., Boyle, A., & Aladjem, D. K. (2005). *Beyond free lunch: Alternative poverty measures in educational research and program evaluation*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Lane, L., & Taber, N. (2012). Experiences with cultural capital in education: Exploring the educational life stories of first-generation postsecondary students. *The Canadian Journal for the Study of Adult Education*, *24*(2), 1–16.

Lareau, A. (1987). Social class differences in family–school relationships: The importance of cultural capital. *Sociology of Education*, *60*(2), 73–85.

Lareau, A. (2001). Linking Bourdieu's concept of capital to the broader field: The case of family–school relationships. In B. J. Biddle (Ed.), *Social class, poverty, and education: Policy and practice* (pp. 77–100). New York, NY: Routledge Falmer.

Lareau, A., & Weininger, E. B. (2003). Cultural capital in educational research: A critical assessment. *Theory and Society*, *32*, 567–606.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity. *Educational Researcher*, *31*(1), 3–12.

Lee, J.-S., & Bowen, N. K. (2006). Parent involvement, cultural capital, and the achievement gap among elementary school children. *American Educational Research Journal*, *43*(2), 193–218.

Martin, M. O., & Kelly, D. L. (Eds.). (1996). *Third International Mathematics and Science Study: Technical report. Vol. 1. Design and development*. Chestnut Hill, MA: Boston College.

Martin, M. O., Mullis, I. V. S., & Kennedy, A. M. (Eds.). (2007). *PIRLS 2006 technical report*. Chestnut Hill, MA: Boston College.

May, H. (2002). *Development and evaluation of an internationally comparable scale of student socioeconomic status using survey data from TIMSS*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia, PA.

McGregor, C. (1997). *Class in Australia: Who says Australia has no class system?* Ringwood, Melbourne, VIC, Australia: Penguin Books.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. New York, NY: The Guilford Press.

Merola, S. S. (2005, August). *The problem of measuring SES on educational assessments*. Paper presented at the 100th Annual Meeting of the American Sociological Association, Philadelphia, PA.

Mueller, C. W., & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development*, *52*, 13–30.

Mullis, I. V. S. (2002). *Background questions in TIMSS and PIRLS: An overview*. Paper commissioned by the National Assessment Governing Board, Washington DC.

Mullis, I. V. S., Kennedy, A. M., Martin, M. O., & Sainsbury, M. (2006). *PIRLS 2006 assessment framework and specifications* (2nd ed.). Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., Olson, J. F., Preuschoff, C., Erberber, E. ... Galia, J. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report: IEA's Progress in International Reading Literacy Study in primary schools in 40 countries*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K., & Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 assessment frameworks*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, MA: Boston College.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *TIMSS 2007 technical report*. Chestnut Hill, MA: Boston College.

Organisation for Economic Co-operation and Development (OECD). (1999). *Measuring student knowledge and skills: A new framework for assessment*. Paris, France: Author.

Organisation for Economic Co-operation and Development (OECD). (2007). *PISA 2006: Science competencies for tomorrow's world. Vol. 1. Analysis*. Paris, France: Author.

Organisation for Economic Co-operation and Development (OECD). (2009). *PISA 2006 technical report*. Paris, France: Author.

Pearce, J., Down, B., & Moore, E. (2008). Social class, identity and the "good" student: Negotiating university culture. *Australian Journal of Education*, *52*(3), 257–271.

Perl, L. J. (1973). Family background, secondary school expenditure, and student ability. *The Journal of Human Resources*, *8*(2), 156–180.

Preuschoff, A. C. (2010). *Using TIMSS and PIRLS to construct global indicators of effective environments for learning*. Unpublished doctoral dissertation, Lynch School of Education, Department of Educational Research, Measurement, and Evaluation, Boston College, Boston, MA.

Priest, S. H. (2009). *Doing media research: An introduction* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage Publications.

Sandoval-Hernández, A. (2012). *An exploratory structural equation modeling approach to evaluate sociological theories in international large-scale assessment studies*. Submitted for publication.

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

Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B. (2010). *ICCS 2009 international report: Civic knowledge, attitudes, and engagement among lower-secondary school students in 38 countries*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *3*(75), 417–453.

Striker, L. J. (1980). "SES" indexes: What do they measure? *Basic and Applied Social Psychology*, *1*(1), 93–101.

Swinnerton, K. A., & Wial, H. (1995). Is job stability declining in the U.S. economy? *Industrial and Labor Relations Review*, *48*(2), 293–304.

Treiman, D. J. (1977). *Occupational prestige in comparative perspectiv*e. New York, NY: Academic Press.

United Nations (UN). (2010). *World population policies 2009*. New York, NY: Author.

United Nations Educational, Scientific, and Cultural Organization (UNESCO). (1997). *International Standard Classification of Education (ISCED 1997)*. Paris, France: Author.

Van Damme, J., Liu, H., Vanhee, L., & Pustjens, H. (2010). Longitudinal studies at the country level as a new approach to educational effectiveness: Explaining change in reading achievement (PIRLS) by change in age, socio-economic status and class size. *Effective Education*, *2*(1), 53–84.

Van de Vijver, F. J. R. (2003a). Bias and equivalence: Cross-cultural perspectives. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 143–155). Hoboken, NJ: John Wiley & Sons.

Van de Vijver, F. J. R. (2003b). Bias and substantive analyses. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 207–234). Hoboken, NJ: John Wiley & Sons.

Vryonides, M. (2007). Social and cultural capital in educational research: Issues of operationalisation and measurement. *British Educational Research Journal*, *33*(6), 867–885.

White, S. B., Reynolds, P. D., Thomas, M. M., & Gitzlaff, N. J. (1993). Socioeconomic status and achievement revisited. *Urban Education*, *28*(3), 328–343.

Wolf, C. (2005). Measuring religious affiliation and religiosity in Europe. In J. H. P. Hoffmeyer-Zlotnik & J. Harkness (Eds.), *Methodological aspects in cross-national research* (pp. 279–294). Manheim, PA: ZUMA.

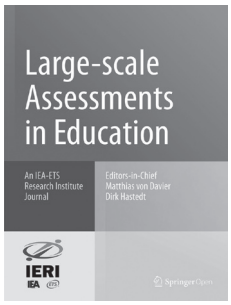
Yang, Y. (2003). Dimensions of socio-economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, *47*(1), 21–41.

Yang, Y., & Gustafsson, J.-E. (2004). Measuring socioeconomic status at individual and collective levels. *Educational Research and Evaluation*, *10*(3), 259–288.

Yang-Hansen, K. (2008). Ten-year trend in SES effects on reading achievement at school and individual levels: A cross-country comparison. *Educational Research and Evaluation*, *14*(6), 521–537.

The IERI Monograph Series is now Large-scale Assessments in Education

(a Springer Open Access Journal)



Large-scale Assessments in Education: An IEA-ETS Research Institute Journal

Editors-in-Chief:

Matthias von Davier, *Educational Testing Service, Princeton, United States*
Dirk Hastedt, *IEA Data Processing and Research Center, Hamburg, Germany*

Managing Editor:

Marta Kostek-Drosihn, *IEA Data Processing and Research Center, Hamburg, Germany*

Journal Scope

Large-scale Assessments in Education is a joint publication of the International Association for the Evaluation of Educational Achievement (IEA) and Educational Testing Service (ETS). The articles in this journal contribute to the science of large-scale assessments, help disseminate state-of-the-art information about empirical research using these databases and make the results available to policy makers and researchers around the world.

Papers suitable for publication in the journal focus on improving the science of large-scale assessments and make use of data collected by programs such as IEA's TIMSS, PIRLS, ICCS, and SITES; U.S.-NAEP; OECD's PISA, PIAAC and TALIS; IALS, ALL and others. If you have questions or concerns about whether your article adheres to the purpose of this journal, please contact the editors directly at editorial@largescaleassessmentsineducation.com. For information on the submission process, please visit www.largescaleassessmentsineducation.com.

Large-scale Assessments in Education is a SpringerOpen journal and features:

- Open access—publications are freely available online
- High-level peer review, production and publishing processes
- Authors retain copyright to their work
- High visibility through different platforms
- Easy compliance with open-access mandates
- Citation tracking and inclusion in bibliographic databases
- Article processing charges are covered by the Institute.

Continuously accepting papers. Please send your submissions to:

www.largescaleassessmentsineducation.com