Towards the development of contextual questionnaires for the PISA for development study

J. Douglas Willms

Lucia Tramonte

WORKING PAPERS

OECD

DIRECTORATE FOR EDUCATION AND SKILLS

TOWARDS THE DEVELOPMENT OF CONTEXTUAL QUESTIONNAIRES FOR THE PISA FOR DEVELOPMENT STUDY

Education Working Paper No.118

by J. Douglas Willms and Lucia Tramonte, University of New Brunswick

J. Douglas Willms, University of New Brunswick (willms@unb.ca)
Michael Ward, Senior Policy Analyst, DCD/GPP (Michael.Ward@oecd.org)

JT03375715

OECD EDUCATION WORKING PAPERS SERIES

TOWARDS THE DEVELOPMENT OF CONTEXTUAL QUESTIONNAIRES FOR THE *PISA FOR DEVELOPMENT* STUDY

Abstract

The aim of this paper is to describe the technical issues to be addressed in enhancing the Programme for International Student Assessment (PISA) contextual questionnaires instruments for the PISA for Development (PfD) study. We discuss the conceptual framework for the contextual questionnaires used in PISA, describe the evolution of the PISA contextual questionnaires, review the measures used in several other international studies, and consider how the PISA data have been used to address the policy questions relevant to the OECD member countries. This research, alongside discussions with key stakeholders, including those from participating countries, enabled us to identify seven themes in which the PISA contextual questionnaires could be enhanced and made more relevant for low- and middle-income countries: early learning opportunities, language at home and at school, family and community support, quality of instruction, learning time, socioeconomic status, and school resources. We discuss various options for enhancing these measures.

Résumé

Ce document a pour objectif de présenter les questions techniques devant être examinées dans le cadre du renforcement des instruments des questionnaires contextuels du Programme international pour le suivi des acquis des élèves (PISA) en vue de l'étude PISA pour le développement (PfD). Nous étudions le cadre conceptuel des questionnaires contextuels utilisés dans l'enquête PISA, décrivons l'évolution de ces questionnaires, examinons les indicateurs utilisés dans plusieurs autres études internationales et analysons la façon dont les données PISA ont été utilisées pour répondre aux questions stratégiques pertinentes dans les pays membres de l'OCDE. Ces recherches, parallèlement aux discussions menées avec les principales parties prenantes, notamment celles des pays participants, nous ont permis de mettre au jour sept grands thèmes pour lesquels les questionnaires contextuels PISA pourraient être renforcés afin d'accroître leur pertinence pour les pays à revenu faible ou intermédiaire : l'accès des jeunes enfants aux possibilités d'apprentissage ; la langue dans les cadres privé et scolaire ; le soutien familial et communautaire ; la qualité de l'instruction ; le temps d'apprentissage ; le statut socio-économique ; et les ressources scolaires. Nous analysons différentes possibilités de renforcement de ces indicateurs.

TABLE OF CONTENTS

Tables

Introduction and background

The Programme for International Student Assessment (PISA) was established in 1997 by member countries of the Organisation for Economic Co-operation and Development (OECD) to provide evidence on students' performance in reading, mathematics and science (OECD, 2001). The surveys included student and school questionnaires which collected contextual data on a range of topics deemed useful for informing policy development within the member countries. The first PISA study, conducted in 2000, included 28 member countries and four partner countries.[1] Another 13 partner countries completed the 2000 assessment in 2002. The 2003 PISA included 30 member countries and 11 partner countries. By 2012, PISA included 34 member countries and 31 partner countries.

The expansion of PISA to include non-member partner countries has presented a number of challenges as these countries tend to have lower levels of income and wealth. The majority of the partner countries that have participated in at least one PISA cycle are upper middle-income countries, with a per capita Gross National Income (GNI) between about USD 4 000 and USD 12 000 in 2011; a further 8 countries are lower middle-income or low-income countries (Bloem, 2013). Many of these countries scored well below the OECD average on the PISA tests, and in many cases a substantial proportion of students scored at or below Level 1 of the PISA proficiency levels. Moreover, many low-scoring students scored at or near the floor of the test, and thus their "true score" may be even lower than their estimated PISA score (Nonoyama-Tarumi and Willms, 2010). In addition, despite the continual development and updating of the contextual questionnaires for every PISA cycle, the student and school contextual questionnaires are not specifically designed to capture the classroom and school contexts of the least economically developed, low- and middle-income countries or to address some of the particular policy issues confronting these countries. This is especially problematic for the measures describing classroom and school resources and the measures comprising the PISA index of economic, social, and cultural status (ESCS). Finally, in the majority of the partner countries there is a larger proportion of the population that is no longer in school at age 15.

The purpose of this paper is to describe the technical and related issues to be addressed in enhancing the PISA contextual questionnaires and data collection instruments for the PISA for Development (PfD) study to make them more relevant to some of the least economically developed countries in the world. The paper examines:

a) existing PISA measures (constructs, indices and variables) and the potential to collect reliable and valid data describing the educational and family contexts in low- and middle-income countries (based on the current PISA frameworks);

b) the analytic requirements to develop meaningful *constructs* and *indices* that can be used cross-nationally while still being based on the PISA frameworks; and

c) the use of these measures in analyses to address policy questions relevant to national and international policy questions.

The focus is on describing the issues, options and possible ways forward for the enhancement of existing PISA instruments, such as by adding to or expanding existing scales, rather than creating new instruments.

---

1. Sub-national economies, such as Shanghai, China and Miranda-Venezuela have participated in PISA, following the same data collection and reporting procedures as member and partner countries. In this paper they are simply referred to as partner countries.

The paper begins with a discussion of the framework for the contextual questionnaires used in PISA and the major policy questions that the data have been used to address them in the OECD member countries. A policy framework developed for a middle-income country is also considered.

The second section describes the evolution of the contextual questionnaires used in PISA, including the main constructs, indices and scales. It also includes a review of the measures used in several other international studies to help inform how the current questionnaires can be enhanced to better fit the needs of a wider range of partner countries.

The third section reviews the PISA ESCS index and other measures of socioeconomic status (SES). The enhancement of the current measures comprising ESCS is not straightforward, as some of the questions required for assessing ESCS in low- and middle-income countries likely differ among partner countries as well as between rural and urban populations within partner countries. The section discusses some options on how to address this challenge. It concludes with a discussion of the potential modification of the PISA measures of school resources.

The fourth and final section discusses the challenges and options for developing the contextual questionnaires.

The preparation of this first version of the paper entailed a review of the literature on contextual questionnaires relevant to large-scale international studies, an analysis of the PISA questionnaires from 2000 to 2012, an analysis of the questionnaires from other international studies, and an examination of the methodological approaches taken to the measurement of socioeconomic status and school resources. The paper has been elaborated further following the discussions with key stakeholders, including those from participating countries at the PISA for Development Technical Workshop: Enhancing PISA's Contextual Questionnaires, which was held in Washington, D.C. from 10-11 April 2014.

1. The role of contextual questionnaires in large-scale international studies

1.1. The PISA questionnaire framework

The questionnaire framework set out for PISA 2012 was designed to include core content that could be kept comparable across cycles (OECD, 2013a, p. 189 ff.). It includes a set of domain-independent measures that can be used in every PISA cycle, as well as domain-specific measures relevant to the target domain of each cycle. The domain-independent measures include:

- Student-level inputs (e.g., grade, gender, family background)
- Classroom instructional processes (e.g., learning time, disciplinary climate, teacher support)
- School-level contexts (e.g., school resources, human resources, community size)
- School-level processes (e.g., assessment and evaluation policies, professional development)
- School-level contexts (e.g., school resources, human resources, community size)
- Non-cognitive outcomes (e.g., truancy, engagement, sense of belonging)

The student and school questionnaires include measures on these themes as well as domain-specific measures, such as instructional quality, motivation, and opportunity to learn relevant to a specific school subject.

Data from the student and school contextual questionnaires are complemented with data from other sources, including teacher and parent questionnaires which are administered in some countries, system-

level data collected by INES, and other questionnaires on particular themes such as interactive communication technologies.

The challenge in enhancing the existing student and school questionnaires for PfD is to include the core PISA content, with individual measures extended where necessary, as well as content relevant to the specific educational policies in the partner countries. In PISA 2012 a rotated design for questionnaire administration was used in order to collect data on a wider range of topics. The decisions about which content to include must also take into account decisions about the form of the outcome assessments for PfD (e.g., a simpler design with less rotation), the merits of rotating questionnaires, and the capacity to analyse and report complex hierarchical data at the country level.

The discussion below considers two approaches to design that can guide these decisions. One is to begin with a model of learning and focus on the factors that can best explain student performance. The other begins with a policy framework and ensures that the contextual questionnaires include the relevant topics to address these questions. The PISA framework includes measures from both of these perspectives and thus serves as an appropriate starting point.

1.2. A model for learning, engagement, and health

The PISA student and contextual questionnaires have to a large extent been based on a long-standing model of learning set forth by Carroll (1963). His model included five factors that contribute to the effectiveness of classroom instruction: aptitude, ability to understand instruction, perseverance, opportunity to learn, and quality of instruction. He distinguished between aptitude – students' general intellectual ability – and ability, which included both their general ability and the prior knowledge required to understand material being taught. The term, perseverance, which refers to the time a student is willing to actively engage in an activity, can be considered part of the broader construct of student engagement. Slavin (1984) adapted Carroll's model and used the term, incentives, to refer to aspects of perseverance that are more directly under the control of teachers.

In his analysis of PISA 2000 data, Willms (2003) argued that student engagement is an important schooling outcome in its own right, which should sit alongside academic performance as a key measure of student success. While most models of learning situate engagement as a predictor of academic achievement, the causal link for many students is predominantly in the other direction, that is, achievement causes engagement. This is especially the case in low- and middle-income countries in which a smaller proportion of the student populations makes the critical transition from "learning-to-read" to "reading-to-learn". Children who do not learn to read fluently with a high level of comprehension by about age 8 or 9 are most often struggling readers through to secondary school (Francis et al., 1996; Kozminsky and Kozminsky, 2001; Morris, Bloodgood and Perney, 2003; O'Reilly and McNamara, 2007). After the primary school years, they are expected to learn subject-matter content in subjects such as social studies, health, history, science, and mathematics. Those who cannot read well fall further and further behind, and by grade 6 or 7 show signs of disengagement. Thus, "engagement begets learning and learning begets engagement" (Willms, 2014). This is the case not only in students' day-to-day classroom experiences, but also from a life-course perspective (Heckman, 2008). The model used by The Learning Bar, which collects student, teacher, and parent contextual data in Canada and Australia from over 500 000 participants every year, maintains that students' physical and mental health is also a key outcome, as important as engagement and achievement, and that achievement, engagement, and health affect each other in an interactive process that begins during the primary grades and continues through to adulthood (Willms, 2014).

Students' performance on the PISA tests at age 15 represents the *cumulative* result on children's experiences at home, in the community, and in school since birth, or even earlier. This fact is often lost in

many policy discussions, as PISA is a survey of students who are attending secondary school. Moreover, in an effort to capture the most salient factors explaining student performance, the PISA context questionnaires have stressed the importance of classroom and school factors that affect student performance in reading, mathematics and science. However, in low- and middle-income countries, especially in rural areas and low socioeconomic communities, children's early family and school experiences likely play a more dominant role than in OECD member countries.

1.3. A policy framework

The second approach to determining content for the student and school contextual questionnaires is to start with a policy framework, or a set of policy questions, and discern what kinds of data are required to address these questions. This approach generally results in a call for a wider range of content than one that follows a Carroll-style learning model that attempts to capture the main determinants of schooling outcomes.

The recently published report, *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices* (Volume IV) (OECD, 2013b), includes a comprehensive set of policy issues that have been addressed by PISA since its inception. They are categorised into four themes:

- selecting and grouping students;

- resources invested in education;

- school governance, assessments and accountability; and

- the quality of learning environments, including school resources, teaching practices, and administrative leadership.

Arguably the biggest challenge facing school systems worldwide is meeting the needs of students with differing ability. Children vary substantially in their language and cognitive skills when they enter school, which affects the likelihood that they will be successful readers by the end of their primary school years. As they proceed through school, the diversity in skill levels increases, reflected in the large variation in reading, mathematics, and science skills observed in PISA at age 15. School systems manage this diversity in various ways, beginning when children enter school. Some of the more prominent strategies are varying the age of entry to school, grade repetition, special schools and school programs, selection into different types of schools, ability grouping within classrooms and within schools, parental choice of schools and school programs, and funding arrangements for private schools or charter schools. The PISA context questionnaires collect data on all of these factors.

Policy questions concerning resources invested in education inquire about the effects of financial, human, material, and time resources. Most of the data on school resources are collected from the school principal or site administrator with the school questionnaire.

Questions about school governance are concerned with how student performance is related to school autonomy, school choice, private and public schooling, principal leadership and parental involvement. The framework also includes questions about the relationship between student performance and various approaches to assessment and accountability.

The last theme of the policy framework pertains to truancy and school climate. The questions on school climate are based on a traditional "school-effects" approach to analysis, with data collected on factors such as teacher-student relations and classroom disciplinary climate.

These four policy themes are certainly relevant to member and partner countries alike. However, when considering the enhancement of the PISA contextual questionnaires to make them more relevant to low- and middle-income countries, we need to consider the relative importance of these themes and how the current scales can be expanded, and finally ask whether there are one or two additional themes, such as the effects of reducing the burden of poverty on students' achievement.

In a paper for UNESCO, "Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems", Willms (2006) showed how socioeconomic gradients and school profiles, which describe the relationship between student performance and the SES of their families at the student and school levels respectively, can play a key role in policy discussions about performance, equality, equity, and effective interventions to raise school performance and reduce inequalities. This framework, which has been used in many of the PISA national and international reports, describes five types of interventions:

- "*Universal interventions* strive to increase the educational performance of all children through reforms that are applied equally across the schooling system.

- *SES-targeted interventions* aim to improve the educational performance of students with low socioeconomic status by providing a specialised curriculum or additional instructional resources.

- *Compensatory interventions* provide additional economic resources to students from low-SES backgrounds.

- *Performance-targeted interventions* provide a specialised curriculum or additional instructional resources for particular students based on their levels of academic performance.

- *Inclusive interventions* strive to include marginalised students into mainstream schools and classrooms."

In the regional report of the Grupo Iberoamericano de PISA, *Iberoamerica in PISA 2006* (OECD, 2010, p. 211), the gradient and school profile schema was extended to include *learning resource graphs*. These graphs displayed, for each country and region, the between-school variation in ten key learning resources measured in PISA that affect student performance. The graphs also included estimates of the "school effect" for each of ten factors, based on a two-level (student and school) hierarchical linear model (HLM) analysis for each county and region. Finally, they provided estimates of the school-level correlations of the learning resources variables and school mean SES.

The analyses in *Learning Divides* and the Iberoamerican report reveal five important findings relevant to enhancing the PISA contextual questionnaires for low and middle-income partner countries:

1. More of the variation in student performance is between schools than within schools, compared with high income countries.

2. The between-school variation in performance is entrenched in the "learning divides" between rural public, urban public, and private schools.

3. Schools' material and human resources vary substantially among schools within each country, are significantly related to school performance, and are strongly correlated with school mean SES.

4. Curriculum coverage and the time allocated to the subject vary substantially among schools within each country, are significantly related to school performance, and are strongly correlated with school mean SES.

5. Teacher quality is very difficult to measure with contextual questionnaires. Consequently, the current measures in most international studies tend to have weak effects on student performance and in some cases the estimated effects are negative.

The Iberoamerican report concluded with four challenges for future assessments:

1. "External assessments need to make use of instruments, which allow us to improve our understanding, and influence student results, in the following areas:

    – classroom and school processes;

    – the functioning of teacher teams, and collective and individual teacher strategies;

    – school organisation, the autonomy of the educational community and shared responsibility in teaching;

    – involvement of those concerned with education, and especially families; and

    – family and student attitudes towards learning.

2. It is essential that progress be made in the [assessment of] added value, which schools and teachers bring to student learning.

3. Measurement instruments must be fine-tuned so that they allow for a more accurate analysis of the most disadvantaged backgrounds, and the most varied social and economic environments.

4. In addition to broadly based policies aimed at making improvements in schools and students, it is crucial to develop specially adapted policies and actions that improve conditions in specific schools. These should also favour the learning needs of students who require certain types of support and guidance."

Willms (2012) set out a framework for linking state and national policy questions to data and analysis and reporting. The schema was developed for consideration of analyses for Prova Brazil, and later modified by The Learning Bar Inc. (www.thelearningbar.com). The framework, which is shown in Table 1, includes seven topics. It differs from the frameworks presented above in that it places greater emphasis on the use of data as "leading indicators" and for assessing equality and equity.

Table 1.   The Learning Bar Policy Framework

Student Performance	What are the current levels of student performance? What are the current levels of student growth? What are the long-term trends in student performance and growth? To what extent do levels of performance and growth vary within and among schools?
Drivers of Performance	To what extent are levels of student performance and student growth related to school processes and practices?
Intervention Forecasting	Can we identify students that require specific types of interventions?
Resource Allocation	To what extent do schools vary in their levels of socioeconomic status? To what extent are low socioeconomic status students concentrated in certain schools?
Equality	To what extent do levels of performance and rates of growth vary among sub-populations, such as socioeconomic groups, immigrant and non-immigrant students, aboriginal and non-aboriginal students, males and females?
Equity	To what extent do students from differing sub-populations have access to key school resources and processes?
Successful Transitions	Are students acquiring the personal assets they need to make successful transitions at key stages of the schooling system and to achieve the learning and career goals that will enable them to become ethical citizens with an entrepreneurial spirit?

Source: www.thelearningbar.com

1.4. Summary

Considering the learning model and policy framework perspectives together, and through consultation with the participating PfD countries, there are at least seven themes in which the PISA contextual questionnaires and data collection instruments could be enhanced and made more relevant for low- and middle-income countries by adding new items, and to a limited extent, by developing new measures. These include:

1. **Early learning opportunities**. The early life-course experiences of many children differ substantially within the partner countries and from those of children in most OECD countries.

2. **Language at home and school**. In several low- and middle-income countries, students' first language differs from the language of instruction at school. Also, in some countries, the language of instruction during the primary grades, when children are learning to read, differs from the language of instruction after the third or fourth grade.

3. **Family and community support**. The nature and extent of family and community support also differs, not only because of cultural differences, but also due to the large number of children living in poverty in many of the partner countries.

4. **Quality of instruction**. Quality of instruction is the most important driver of student achievement, but also the most difficult to define and measure. The nature of instruction in partner countries differs from that of most OECD countries, especially in rural schools.

5. **Learning time**. Learning time in low- and lower-middle income countries differs from that of high-income countries in several ways. In many cases children start school later, they miss many days of school during the primary school period, their levels of school attendance are generally lower, and they are more likely to repeat grades. Many children work in part-time jobs outside

the home from an early age. Moreover, there appears to be considerable variation in class time devoted to the three core PISA subjects and the curriculum coverage is not as deep.

6. **Socioeconomic Status**. The measure of SES, or what is called ESCS in PISA, needs to be modified to capture lower levels of education and lower levels of income and wealth than the current PISA scales. This can feasibly be accomplished by supplementing the lower end of the scales with additional items.

7. **School Resources**. The measure of school material resources needs to encompass very low levels of school resources, including aspects of infrastructure that are often taken for granted in high-income countries. Also, the measure needs to make a clear distinction between material resources, schooling processes, teachers' working conditions and human resources.

The challenge will be in adapting the contextual questionnaires without extending questionnaire time, as countries could easily argue that *all* of the content of the current PISA questionnaires is also relevant to their contexts. The changes need to take into account the strategy used to collect performance data in PfD; if all three domains are assessed equally then the domain-specific content will need to be modified.

Also, many of the child and family issues raised could be addressed with a parent questionnaire, which is offered as a country option in PISA. The increased costs may be warranted if this is the best way to collect valid contextual data in low- and middle-income countries. Similarly, many of the classroom and school constructs could be better addressed with a teacher questionnaire. In this paper, our emphasis is on enhancing the current student and school questionnaires, but we do not rule out the possibility of parent or teacher questionnaires.

The next section of the paper summarises the content of the PISA questionnaires and the questionnaires used in several other international studies.

2. Contextual questionnaires in PISA and other international studies

2.1. The evolution of measures in PISA, 2000-2012

The PISA student and school contextual questionnaires have evolved considerably since the first PISA survey in 2000 (Klieme et al., 2014) The majority of measures that have been used across all five cycles pertain mainly to demographic characteristics of students and their families (e.g., grade, gender, parental education) and of schools (e.g., type of school, enrolment). We use the term "measures" to refer to single-item measures such as grade and gender, and multi-item constructs such as teacher-student relations. Annex A summarises the coverage of the questionnaires over the past five PISA cycles. The measures are grouped into four categories: child, family, classroom and school. For each measure, S and A are used to distinguish between the student and school questionnaires respectively. The number of items for the measures is shown in brackets. For example, for "values schooling outcomes", S (4) indicates that the construct is measured with four items in the student questionnaire. The final column indicates the number of PISA cycles in which the measure was used.

We distinguish between child and family variables as this separation is relevant for conceptual reasons when one considers the validity of student *versus* parent-reported data. However, in PISA, from an analytic perspective the family measures are normally attached to each child's record and treated as child-level variables. The distinction between classroom and school factors is especially important, though, as there is a strong literature indicating that classroom effects on student performance are stronger than school effects (Hill and Rowe, 1996; Mortimore, Sammons, Stoll, Lewis and Ecob, 1988; Scheerens, Vermeulen and Pelgrum, 1989; Willms, 2010).

2.1.1. Core demographic measures

Measures that were included in at least four of the five cycles are indicated with bold text in Table 2. We consider these "core content" in the sense that they are highly likely to be included in the PfD contextual questionnaires. However, a few of them will need to be modified to fit the context of the partner countries.

Six child demographic measures were included in at least four cycles of PISA: date of birth, age, gender, country born, age when immigrated, and school program. The questions for these measures are not likely to require much modification for PfD. The exception is "school program", which is adapted to fit each national context. Eleven family measures were included in at least four cycles of PISA: family structure, mother's occupation, father's occupation, mother's education, father's education, home possessions, educational possessions, books in the home, the country in which the mother was born, the country in which the father was born, and the language spoken at home. Several of these measures will be used to develop a composite measure of socioeconomic status, which is discussed in a later section of this report. Other demographic factors, such as those associated with minority ethnic or religious cultures will need to be addressed with national (optional) questions.

Ten measures have been consistently used to describe the demographic characteristics of the school: class size, type of community, enrolment, school type, grades covered, funding sources, selective school, infrastructure, computing resources, and teacher full-time equivalents. At least one of these measures – school infrastructure – will need to be enhanced for PfD.

Two measures of student engagement have been used in at least four cycles of PISA: values schooling outcomes and attitudes towards the subject. Two other measures of student engagement, attendance and sense of belonging at school, were used in three of the five PISA cycles. These two measures were used to construct the two principal variables for the OECD report on student engagement (Willms, 2003).

In addition, six measures have been used consistently in PISA to describe the classroom and school context: teacher-student relations, classroom disciplinary climate, use of assessment, barriers to student learning, streaming, and autonomy.

2.2. Constructs measured in other international studies

Annex B identifies several measures used in other large-scale international studies that could inform the enhancement of the PfD contextual questionnaires. They are described briefly in Annex C. The studies selected include those that include a significant number of the PISA partner countries or emphasise data collection in low- and middle-income countries. The table includes all of the measures in Annex B, with the first column repeating the extent of usage in PISA. The potentially "new" variables are shown with slightly darker shading. While these other international studies include a wide array of other variables, we selected those that could potentially contribute to the seven themes identified in the previous section where the PISA contextual questionnaires could be enhanced and made more relevant for low- and middle-income countries. The theme is identified in the first column with the following abbreviations: early learning opportunities (ELO), language at home and school (LHS), family and community support (FS), quality of instruction (QI), learning time (LT), school resources (RES), and socioeconomic status (SES).

2.2.1. Early learning opportunities

The table shows that there are only two measures on early learning opportunities: school readiness and reading to the child.

2.2.2. Language at home and school

The PISA student questionnaires have included in every cycle a question about the language spoken at home. Other surveys have included questions about other languages the student speaks and the student's familiarity with the language of the test. Little attention has been paid to the broader question of the structural features of the school system regarding the language of instruction during the primary school period and the language of instruction after grade 3 or 4.

2.2.3. Family and community support

The PISA questionnaires have included measures of family support, but only in certain cycles. These include measures of social and cultural communication with parents, cultural capital, and family involvement. The other international questionnaires include measures of social capital, parental disciplinary practices, and a general measure of communication and involvement of the family.

In an analysis of the 2000 PISA data, we examined the effects of cultural capital on students' reading performance (Tramonte and Willms, 2010). We described two types of cultural capital: one that is static, associated with the "highbrow" activities and practices of parents, and one that is relational, pertaining to the cultural interactions and communication between children and their parents. The results showed that relational cultural capital had stronger effects on children's reading performance. However, our analysis was limited to OECD member countries and we can only speculate on how various aspects of family and community involvement and support affect student performance in low- and middle-income countries. We suspect that static forms of cultural capital play a strong role in determining whether a child attends a public or private school, as would other variables correlated with SES. However, once the type of school is determined, relational forms of cultural capital would explain some of the within-school variation.

One way to conceive of community support is to consider it an aggregate of parent involvement and support; schools with strong community support are those that have a strong parent community that is involved in their children's learning and supporting teachers in their day-to-day work. This could also include volunteering at school and participating in the governance of the school. One can also consider a broader form of community support that includes community members other than parents, who may support the school financially or be involved as consultants or advisory board members. In some countries, there are different types of schools which grew out of community "grass roots" organisations, or through alliances with international agencies and NGOs. Some of these schools are considered stand-alone "community schools" providing an alternate form of education than that offered by the state, while others are schools that intend to become part of the state system (Rose, 2003).

2.2.4. Quality of instruction

The PISA framework has consistently included questions about classroom context, such as the classroom disciplinary climate and teacher-student relations. In 2009 it was extended to include questions about the nature of classroom instruction, in particular, questions assessing time spent on activities that would be considered part of a direct instruction model and questions about "cognitive activation" (Klieme et al., 2009). Other cross-national studies, notably PIRLS, LLECE, PASEC and SACMEQ, have included items pertaining to more basic instructional activities, which may be useful for extending the current framework.

2.2.5. Learning time

The construct of learning time for low- and middle-income countries needs to capture learning time in and out of school. In other studies, in-school learning time is measured with teacher questionnaires in two ways: by computing the average duration of classes, school year, and length of statutory holidays; and by

the extent to which the curriculum of a specific discipline is covered in class. The out-of-school learning time is measured indirectly with questions about students' interest and involvement in reading practices.

2.2.6. Socioeconomic status

The other international studies include a number of questions that can be considered for enhancing the PISA measure of socioeconomic status. Two dominant themes pertain to the child's experiences out of school: participation in the labour force (e.g., child labour, hazardous labour conditions, and hours of labour per week); and health and well-being, including nutrition. Students are also asked about the time and means of travel to reach school and where they live during the school week. At the family level, two themes emerge: the literacy practices of the parents and home possessions. The list of home possessions includes items that are more relevant to low- and middle-income countries (e.g., properties of the dwelling, electricity, and water and sanitation).

2.2.7. School resources

The questions used in other international studies can be grouped into four themes: material resources, instructional processes, teachers' working conditions, and teachers' human resources. The measures of material resources do not add substantially to those already included in PISA, and therefore we will likely need to look to some of the partner countries' questionnaires to identify some of the key resources for low income schools. The other questionnaires tend to place considerable emphasis on teachers' classroom practices and activities (e.g., goal setting, differentiation, active learning); however, there is no unifying framework. A number of questions pertain to teachers' working conditions, such as their job satisfaction, economic incentives, and the extent of collaboration and support. Finally, there are several common measures of teaching human resources, such as teachers' training and years of experience.

The work associated with these seven themes will be complemented with two other expert papers commissioned for the development of the PfD. The key challenge is in achieving international comparability with the current PISA measures and scales, without creating entirely new instruments that cannot easily be linked to the PISA framework.

2.3. Summary and recommendations

This chapter identified seven themes that are core to enhancing the PfD context questionnaires. The chapter was modified based on suggestions from participating countries during the PfD technical workshop. The suggestions for enhancement and modification were considered in the light of two main criteria: international comparability and integration with the current PISA frameworks.

To address issues regarding the first five themes – early learning opportunities, language at home and school, family and community support, quality of instruction, and learning time – the PISA contextual questionnaires can be enhanced with relatively minor modification and extensions of the existing instruments.

The questionnaires could include one or two questions regarding students' early learning experiences, such as whether they attended a day-care. Data on grade repetition during the early years would also be valuable. The measurement challenge will be whether 15-year-old students can provide valid retrospective data.

The student questionnaire could also include a question regarding students' familiarity with the language of the test. Students may be able to recall the language of instruction when they learned to read in primary school, but any sort of diary or recall question may be challenging for many students. Thus, the

broader question concerning the language of the instruction at varying grade levels would be best addressed with a system level questionnaire.

The PISA measures on family support could be enhanced so that there are clear-cut measures of parental involvement, social capital and cultural capital.

These measures could be supplemented with a measure that focused directly on the role of other community members. Bray (2000) considered degrees of participation in education by communities in various educational functions, such as designing policy, teacher training, and building and maintenance. A simple and straightforward approach to this issue could be an extension of the current PISA School Questionnaire item on who has responsibility for various tasks (e.g., selecting teachers, firing teachers, establishing teachers' starting salaries, etc.) to include other community members alongside the principal, teachers, parents, etc. The measure also needs to incorporate issues concerning security and school safety. Finally, a question on types of community schools could be included in the School Questionnaire.

Consideration should be given to which of the current PISA questions are relevant for the partner countries. The revised questionnaire needs to capture, at a very basic level, the extent to which class time is spent in independent activities, such as working in workbooks, versus small group activity and whole-class teacher-centred instruction. The significant challenge for collecting data on quality of instruction is that it needs to be subject-specific, and this would require survey rotation or a significantly longer survey.

Learning time could be assessed more thoroughly with some measures of how students spend their time in and out of school. The measure of school attendance could also be enhanced. Students' participation in the labour market *also* needs to be captured.

The last two themes – socioeconomic status and school resources – require more extensive modification for PfD. These themes are discussed in the next two sections.

3. The measurement of socioeconomic status

3.1. A definition of socioeconomic status

The term "socioeconomic status" (SES) generally refers to the relative position of a family or individual on an hierarchical social structure, based on their access to, or control over, wealth, prestige, and power (Mueller and Parcel, 1981). Numerous studies have shown that a person's position on an SES hierarchy is related to a wide range of outcomes pertaining to their physical, economic, and social well-being. SES affects people's educational opportunities, their access to certain labour markets, their exposure to health and crime risks, and their lifetime earnings.

Coleman (1988) and Bourdieu (1977) used the economic term, *capital*, as a metaphor for the cultural and social assets that families possess which lead to higher levels of physical, economic, and social well-being. With respect to educational outcomes, Coleman stressed the important role of social networks, reciprocity, trust, and the development of social norms. Bourdieu's concept of cultural capital encompasses a wide range of attributes associated with status, including, for example, certain manners, linguistic styles, dress, and preferences for "highbrow" cultural activities. Schools are middle-class institutions with particular linguistic structures, authority patterns, and types of curricula. Parents that possess considerable cultural capital are familiar with these values and structures and transmit them to their children. Thus, when their children enter school they know intuitively the strategies for negotiating a successful school pathway and these are reinforced by ongoing experiences at home (Lareau, 1987; Lareau and Weininger, 2003). The key roles of social and economic capital in the education process have led to definitions of SES that include these concepts. For example, an expert panel considering the measurement of SES for the

United States National Assessment of Educational Progress, defined SES as "one's access to financial, social, cultural, and human capital resources" (National Centre for Education Statistics, 2012).

In most studies of the effects of families, schools, and communities on children's academic and social-emotional development, SES is operationally defined with measures describing the occupational prestige, educational levels, and the income of the children's parents. Because data on family income are usually difficult to obtain, data on material and educational possessions in the home are used as a proxy for income. The PISA ESCS index is comprised of the occupational status of the parent with the highest occupational status, the number of years of education of the parent with the highest educational attainment, and an index of home possessions that includes measures of wealth, educational possessions, and cultural possessions, including the number of books at home.

3.2. The role of SES in educational studies

Some analysts argue that SES is a useless concept for policy purposes, as there are no policy instruments that simultaneously act upon income, education, and social class (Deaton, 2002). Given that governments need to make decisions about expenditures on education, reduction of poverty, and policies to stimulate employment, it makes sense to examine separately the effects of family income, parental education, and employment status. Also, in analyses aimed at predicting educational performance, an examination of the separate effects of the constituent components of SES is more likely to uncover underlying causal mechanisms.

However, an understanding of the SES-outcome relationship for various outcomes, across jurisdictions and over time, is extremely useful for understanding the key relationships at a macro-level, especially those relevant to issues concerning equality and equity. For example, when comparing the relationships between academic achievement and family background across countries, the estimation of the outcome-SES relationship is useful in understanding the structure of the school system (Willms, 2006; 2010). If one tries to estimate the relationships with each of the constituent components of SES, one cannot easily discern the macro-relationships, in part because the three main components are correlated to varying degrees within each country.

When we consider how best to measure SES, it is useful to consider how it is currently being used in PISA, in the international and national reports and by researchers doing secondary analyses. SES is being used in the following ways:

3.2.1. Estimating SES gradients and comparing them across countries

The gradient is simply the relationship between a child's developmental outcome and SES. Often the research questions pertain to the social determinants of gradients; for example, why are inequalities along social class lines greater in some countries or some communities than in others?

3.2.2. Controlling for family background

Analysts often wish to compare two groups on an outcome measure, but the groups differ in their family background characteristics. SES is considered a confounding factor and is used as a covariate to control for group differences.

3.2.3. Increasing precision of estimates

In some cases, even though groups do not differ substantially in their family backgrounds, the use of SES as a control measure increases statistical power and yields more precise estimates of group differences and of the effects of other variables. Sometimes an analyst wishes to obtain estimates of group differences,

or the effects of other variables, with and without adjustment for SES. For example, in studying the effects of immigrant status on student performance, one would want estimates of the distribution of the outcome measure for immigrant and non-immigrant students, both before *and* after controlling for SES.

3.2.4. Examining interactive effects

SES is useful for examining interactive effects between family background factors and a treatment variable, or between family background factors and the child's sex or ethnicity. For example, a treatment may be effective for low-SES children, but not for high-SES children, or *vice-versa.* Similarly, differences between females and males in their educational achievement can vary, depending on the student's SES. The important relationships can be conveyed more succinctly with a summary measure of SES.

3.2.5. Summarising the distribution of family background

SES is often useful as a summary descriptive variable to denote the distribution of background variables for some aggregate unit such as a school, region, or country. For example, one can use a summary measure of SES to categorise schools as low-SES or high-SES or to describe the heterogeneity of schools' student intake.

3.2.6. Estimating the extent of inclusion or segregation

SES can be used to describe the extent of inclusion or segregation among schools within a region or country (Willms and Paterson, 1995). The PISA measure of ESCS was used in this way in the 2011 *Education at a Glance* (OECD).

3.2.7. Increasing generalizability

A measure of SES is useful for assessing the extent to which the findings of a study are generalizable to some larger population.

3.3. Four key methodological issues

3.3.1. Reflective versus formative approaches

Edwards and Bagozzi (2000) described two different models of the nature and direction of the causal relationship between constructs and measures. A reflective approach assumes that an underlying construct, such as self-esteem, causes various observable indicators, such as how students respond to various Likert questions pertaining to their feelings of self-worth (Bollen, 2002). A formative approach assumes that the observable indicators cause or affect the latent variable. For example, students' sense of belonging at school, their participation in sports and clubs, and their effort and motivation affect their engagement at school.

The approach used determines not only which variables are included in the model; it also affects how they are weighted. Most constructs used in the social sciences, including SES, are based on a reflective approach. With this approach, one can use techniques such as factor analysis to determine the weights, and use psychometric approaches such as Cronbach's alpha to assess reliability. This is the approach taken with the PISA measure of ESCS. A formative approach is "purpose-built" in the sense that the indicators are presumed to cause a latent construct which in turn is related to a particular outcome. For example, we could use multilevel regression to choose indicators of SES that were strongly related to PISA math performance, and base the weighting on the estimates of the regression parameters. Those advocating this approach maintain that SES itself derives its meaning from its relationship to valued outcomes. However, others argue that the formative approach is circular: if one chooses variables that predict the outcome and

then use the construct to describe its relationship to the outcome, then little is gained by having the SES measure (National Center for Education Statistics, 2012).

3.3.2. A tipping point

Many of the relationships between educational and health outcomes and income appear to be fairly linear across a wide range of income levels (Adler et al., 1994). The classic Whitehall study by Marmot and his colleagues found that the health status of British civil servants was related to job status, even at the top end of the status hierarchy. However, there is increasing evidence that health outcomes are not linearly related to SES – the effects diminish at higher levels of the social hierarchy (Boyle and Willms, 1999; Wolfson et al., 1999). Similarly, the relationship between life expectancy and Gross National Product (GNP) at the level of countries is also curvilinear, with diminishing returns at levels above USD 5 000 (World Bank, 1993). Willms (2006) tested this "hypothesis of diminishing returns" with data from PIRLS and PISA. He found that the gradients were linear in most countries, and in countries where they were curvilinear, there was diminishing returns to SES in some countries and increasing returns to SES in others. Willms and Somers (2001) found that the SES gradients for children's reading and mathematics scores were curvilinear in several Latin American countries. However, the curve increased with increasing levels of parental education. They suggested that there may be some minimum level of parental education necessary for children to benefit from elementary schooling in low- and middle-income countries.

The idea that there is a "tipping point" at which children can maximally benefit from rising levels of SES is critical to the current issues pertaining to the enhancement of the PISA contextual questionnaires. If one adopts a formative approach to the choice and weighting of variables in the SES measure, then it may be that the indicators affecting schooling outcomes at very low levels of SES differ from the indicators affecting schooling outcomes at higher levels of SES. One might also hypothesise that in the least economically developed countries the SES indicators that affect performance at the lower end of the scale are strongly predictive of whether children make the critical transition from learning-to-read to reading-to-learn. This opens up the possibility of having two distinct measures of SES, one that includes indicators relevant to children who are living in poverty, and another that includes indicator for children whose SES is low to moderate, but above the "tipping point".

3.3.3. Measurement invariance

Ideally one would like to have a single measure of SES that was consistent across sub-populations – for children of families in rural and urban settings, for immigrant and non-immigrant students, across low-, middle- and high-income countries, and so on. This would facilitate interpretation and reporting, and enable the analyst to readily make cross-national comparisons. Wu, Li and Zumbo (2007) describe measurement invariance: "An observed score is said to be measurement invariant if a person's probability of an observed score does not depend on his/her group membership, conditional on the true score. That is, respondents from different groups, but with the same true score, will have the same observed score." This is virtually impossible, especially if one adopts a formative approach. For example, having a parent with a high school diploma has quite a different meaning in a low- or middle-income country than in a high-income country. Moreover, when considering home possessions as proxies for income or wealth, certain possessions such as owning a bicycle are not invariant, even within the same country. Indeed, with respect to the enhancement of the SES measure for PISA, the measurement invariance associated with living in poverty *versus* not living in poverty far outweighs the issue of measurement invariance across countries.

3.3.4. Other levels of measurement

One could make a case for including certain child, school, or community indicators in a measure of SES. For example, children's participation in the labour force affects family income, is related to other

indicators of SES, and affects their schooling outcomes. Similarly, measures describing the resources available to a student in the school and community are consistent with a contemporary definition of SES – one's access to financial, social, cultural, and human capital resources. However, we believe that including indicators measured at other levels unnecessarily complicates the development of a measure of family SES. Moreover, these variables can be included in the contextual questionnaires alongside a measure of SES.

3.4. An integrated approach for PfD

3.4.1. Extending the current indictors of PISA ESCS to include items at the lower end of the SES scale

One approach to enhancing the PISA measure of SES is to include items that extend the home and educational possession scale such that it captures the level of SES of children living in poverty. In their analysis of Prova Brazil data, for example, Willms, Tramonte, Duarte and Bos (2011) attempted to use items on home possessions to set a cut-point for children living in poverty. However, their Item Response Theory (IRT) analysis revealed that most of home possession items were at the high end of the SES scale, such that the "information" (in an IRT sense) for accurately identifying those living in poverty was limited. More low-SES items are necessary for a more accurate description of the population's SES. The same argument would apply to PISA for low- and middle-income countries. Thus, one could use data from other studies, such as SACMEQ or MICS, to identify items that can be used to extend the scale at the lower end. One could also attempt to develop finer-grained measures of parental education and occupational status.

3.4.2. Developing new "poverty-related" measures

Another approach is based on the notion that there is a tipping point at the lower end of the PISA scale at which the current PISA ESCS lacks validity as a formative measure of SES. If so, there is a need to develop one or more new indicators of SES that apply to children living in poor families. This is different conceptually than simply extending an existing scale at the lower end in that there is a fundamentally different construct for students who are living in poverty. For example, items describing whether there is a radio or TV in the home may suffice to extend the current PISA scale at the lower end, but other items, such as having electricity or running water in the home may be indicators of poverty that are associated with a child's developmental experiences since birth. Similarly, the PISA measure of parental education can be extended lower by including some finer-grained measures of level of education that would situate parents more precisely at the middle or lower primary levels. However, a measure that gauged the extent to which parents were literate and engaged in reading activities at home is a different, albeit related, construct to level of education. The same argument would apply to items describing lower levels of occupational status versus a measure of the extent to which the parents were participating in the labour force.

3.4.3. Combining the two approaches

At this stage, one can pursue the first strategy and attempt to extend the current scales for possessions, parental education, and occupational prestige. The limits to this approach will become evident if *(a)* the resulting scale becomes multidimensional or *(b)* if the new items differ in their relationship to achievement for low- and high-SES students. If so, this would obviate the need to separate indicators that described levels of family poverty.

All of the countries participating in the PISA for Development Technical Workshop, held in Washington, D.C., agreed that a combined approach is required. However, this does not have to violate the two main criteria – international comparability and maintaining a link to the current PISA framework. For example, the PfD questionnaires might include the three current measures of ESCS with new items that extend their range, plus one or more poverty-related measures that sit alongside the three core measures. In the current framework, all countries would complete the three traditional (extended) measures and these

would be used to construct ESCS. The ESCS measure would be used to estimate socioeconomic gradients in the same way as in previous cycles. In addition, some countries would choose to include the new "poverty" measures as a national option. These measures would be used alongside SES in various analyses, as described in section 3.2 above. Their inclusion in PfD would enable some partner countries to get a better purchase on the extent and depth of poverty and how it affects educational outcomes.

The development of a separate set of poverty-related items to extend the current set of PISA items on home possessions presents at least two challenges. One is that it may be impossible to identify possessions that function similarly across partner countries. For example, with supplementary analyses of data from MICS, we found that owning a bicycle has a different position on the socioeconomic scale from one country to the next; that is, there is Differential Item Functioning (DIF). We are somewhat optimistic that one could identify five or six items to form a cross-national scale and it is possible to examine the psychometric properties of several poverty-related items using data such as MICS in advance of the PfD pilot. The set of core items could then be supplemented with national "option" items that have strong local relevance and cultural validity. With this approach it would be possible to have a small "poverty-related" scale that serves the purposes of international comparability and a local SES scale for within-country analyses.

Another challenge is that when one asks students questions about poverty-related items they may be reluctant to give valid answers. These items would be different than items describing educational possessions in the home, which have more face validity in a student questionnaire.

The collection of data on parents' literacy skills and their records of employment would be best collected directly from the parents. Moreover, one cannot simply modify the existing PISA parent questionnaire to capture these data. Rather, it would require an interview approach similar to that used in LAMP and other studies conducted in low-income countries. This would undoubtedly be valuable data but may be too expensive to include as a core feature of PfD. If conducting a parent survey is considered to be too expensive or beyond the scope of PISA, one could consider adding a small number of items to the student questionnaire which would likely improve the measurement of SES. These would include an item on parents' current employment status and a question on parents' literacy practices at home, similar to those used in the International Adult Literacy Survey.

The choice between a reflective versus formative approach to identifying relevant items does not need to be an either/or decision. Research can be conducted using the current PISA data, as well as data from other studies, following a formative approach; that is, choosing items with an eye towards their association with student performance in reading, mathematics, and science. The analyses might also examine the relationship between new poverty items with a wider range of health and education outcomes, based on analyses of data from the other international studies discussed above. This would help address the circularity criticism associated with a formative approach. However, the construction of the ESCS and its underlying scales can use reflective approaches – item response theory and factor analysis for example – which have been a long-standing strength of the PISA framework.

4. The measurement of school resources

4.1. A definition of school resources

The measurement of school resources presents a similar challenge to that of SES, and we could consider the same three potential solutions presented above. In this case, however, we know relatively little about the nature of the relationships between school resources and student performance at lower levels of school resources. We also maintain that there needs to be a clear distinction between material resources, classroom and school processes, teachers' working conditions, and human resources.

In their research on the effects of school resources based on LLECE data, Murillo and Román (2011) helped to clarify the definition of school resources by distinguishing between basic services, didactic facilities, and didactic resources:

4.1.1. Basic services

Basic services at the school include factors such as potable water, sewage services, bathrooms, electricity, and telephones.

4.1.2. Didactic facilities

Most schools have one or more places, other than the classroom, for teaching and learning. These include for example, school libraries, gymnasiums, art and music rooms, science laboratories, computer rooms, and sports fields.

4.1.3. Didactic resources

Didactic resources can include very basic materials such as textbooks and blackboards, through to computers in the school, laptop computers for students and teachers, and quality books in the library.

For the purposes of examining potential enhancements of the PISA measures for PfD, we will use this framework. Other aspects of resources, such as classroom and school processes, teachers' working conditions, and human resources are included as separate themes in the PISA framework.

4.2. The PISA measures on school resources

The PISA 2012 school questionnaire included two sets of questions relevant to school resources. One set collects data pertaining to the ratio of students per computer in the school. The other set examines the extent to which the school administrator feels that the school's capacity to provide instruction is hindered by particular issues. The set includes 12 factors, and countries can add their own factors as a national option. The 12 factors include two items regarding basic services (adequacy of school buildings and grounds, heating/cooling and lighting systems) and one general item regarding didactic facilities (classroom instructional space). Six items pertain to didactic resources (science laboratory equipment, instructional materials, computers for instruction, Internet connectivity, software for instruction, and library materials). Four other items assess whether instruction is hindered by a lack of qualified teachers in the test language, mathematics, science and other subject areas. The school questionnaire also includes a question about how much schoolwork requires students to access the Internet.

4.3. A framework for the measurement of school resources for PfD

Several studies have suggested that school resources do not have a strong impact on student achievement (e.g., see reviews by Sammons, Hillman and Mortimore, 1995; Scheerens and Bosker, 1997). However, the two key studies based on LLECE data (Murillo and Román (2011); Willms and Somers, 2001) suggest that school resources do have substantial effects, even after taking into account the socioeconomic characteristics of students. At least four issues plague this vein of research. The first issue is that school infrastructure and resources may have large indirect effects but small direct effects on student outcomes. For example, qualified and talented teachers are more likely to be attracted to and retained in schools with higher levels of resources. A related issue is that of collinearity. Factors such as the presence of or size of a school library is correlated with school mean SES and whether it is a public or private school. The LLECE studies attempted to control statistically for school mean SES and private schooling, but the accuracy of the findings depends critically on model specification. This can lead to an erroneous conclusion that libraries and school computers are the most important school resources. A third issue is that

it is very difficult to conduct a randomised controlled study on school resources. This would provide more definitive results on the effects of resources. Finally, in the case of didactic resources, their availability does not necessarily mean they are used widely and appropriately by students and teachers.

We maintain that a small set of straightforward questions about basic services, didactic facilities, and didactic resources would be a useful addition to PfD. Some partner countries collect data on school resources as part of their monitoring programs and have developed indices of school infrastructure. PfD could take stock of the various approaches used by partner countries as a starting point to develop a small set of new questions on school resources would enable partner countries to make cross-national comparisons, examine their correlation with student achievement, assess equity of educational provision, and track progress from cycle to cycle. A number of the international studies, such as the LLECE, and local monitoring systems, such as Prova Brasil, include questions which essentially take an inventory of school resources. Analyses conducted prior to the PfD pilot study would enable the research team to develop a coherent assessment framework for school resources.

5. Summary of issues, challenges and options

This section provides a summary of the issues discussed in this report and sets out some of the challenges and potential options for consideration by the countries participating in the PfD pilot study.

5.1. Types of contextual questionnaires

The PISA contextual questionnaires have traditionally included only student and school questionnaires for all participating countries. Logically one should consider what data partner countries wish to collect and then discern whether these data could be best collected from students, parents, teachers, or school administrators. The PISA framework includes student, parent, teacher and school questionnaires which can be adapted for PfD. The main challenge is discerning the best informant for measuring the constructs in the seven core themes.

5.2. Policy framework

The first section set out two general approaches for determining the content of the enhanced questionnaires. One approach is based on a learning model, which is similar to the Carroll (1963) model, which underlies many international studies. The other is based on a policy framework, and three possible frameworks were presented. The PISA framework set out in PISA 2012 is very comprehensive. Nevertheless, a first point of discussion is whether the framework includes all of the topics relevant to PfD countries, or if there should be lesser of greater emphasis on certain issues. For example, one of the needs expressed by the Grupo Iberoamericano de PISA (2010) was that "it is crucial to develop specially adapted policies and actions that improve conditions in specific schools. These should also favour the learning needs of students who require certain types of support and guidance" (p. 211).

5.3. Core themes for enhancement

Our analysis of the constructs and measures used in PISA and other international studies identified seven core areas to be considered for enhancement. These were:

- early learning opportunities;

- language at home and school;

- family and community support;

- quality of instruction;

- learning time;

- socioeconomic status; and

- school resources.

Our analysis suggested that for the first five core areas the necessary enhancements could be achieved with minor modifications of the current PISA measures and the development of a few new constructs, in part based on questions used in other surveys. The measurement of socioeconomic status and school resources will require more extensive analysis and some pilot testing.

A starting point for this work, prior to a pilot study in the partner countries, would be a rigorous and systematic analysis of the student and school items that have been used in the PISA questionnaires. The analysis would examine the extent of inter-country and urban-rural Differential Item Functioning and the relationship between students' academic performance and individual items. The same analyses would also be conducted to examine DIF and the relationships with student performance with potential indices, such as indices of learning time, quality instruction, and school infrastructure.

A second point of departure would be to examine the validity and cross-national comparability of items and indices that have been used in other studies, such as SACMEQ, PASEC, and LLECE. The findings from this work can be enhanced by the data collected from other kinds of studies, such as LAMP and STEP, as well as data collected at the system level by UNESCO.

5.4. Measurement of socioeconomic status

The current PISA measure of socioeconomic status, ESCS, does not include a sufficient number of items at the lower end of the scale to adequately describe the populations of students in the majority of partner countries. We presented two options for addressing this issue:

- extending the current indictors of PISA ESCS to include items at the lower end of the SES scale; and

- developing new "poverty-related" measures.

We have recommended pursuing the first option as a starting point, but suspect that it will have limitations in that the resulting scale will not be uni-dimensional and the new items will differ in their relationship to achievement for low- and high-SES students. Therefore, a combined approach can be pursued, with attention to the goals of international comparability and maintaining a link to the current PISA framework.

5.5. Measurement of school resources

The current PISA questions on school resources focus on administrators' perceptions of whether a lack of school resources hinders the school's capacity to provide instruction. The questions focus mainly on didactic resources. A key question for the PfD participating countries is whether a small set of new measures on school resources that provided data on basic services, didactic facilities and didactic materials is of high priority. This discussion needs to take into account the length of the questionnaire and other competing priorities.

5.6. Integration with main PISA study

The integration of the enhanced measures into the main PISA study presents two challenges. One is simply a matter of questionnaire length. The new content for the student questionnaire may require 10 to 15 minutes of extra content, for example. If so, a comparable amount of content from the main survey will need to be deleted. The second challenge, which is related to the first, is that the current PISA questionnaires have appropriate content for many of the students in schools in the partner countries.

The strategy for addressing these issues depends on decisions about the approach to the measurement of student performance and whether there will be rotated versions of the questionnaires in PfD. The discussion at the technical workshop must also consider the technical capacity of participating countries to analyses data stemming from a more complicated design.

The measures and indices developed for the PfD questionnaires will be designed and implemented in accordance with PISA Technical Standards.

5.7. Trialling the enhanced questionnaires

The development of the measures of SES and school resources will require a number of psychometric analyses, including analyses in an Item Response Theory (IRT) framework. At a minimum we recommend conducting pilot testing of the revised questionnaire in six or seven countries, with 20 schools sampled within each country. The sample should be a strategic sample, with a range of school types (e.g., rural/urban, public/private) sampled within each country. Within schools, we recommend that all 15-year-olds complete the survey, as the costs of collecting and processing population data may not be greater than those associated with within-school sampling.

During the 1st International Advisory Group meeting on PISA for Development, several questions will need to be addressed concerning the practical issues for the pilot study. These include: the review of instruments, challenges for pilot testing (e.g., sampling of schools, obtaining reliable and valid data from school administrators), and the integration of content into the current PISA framework.

REFERENCES

Adler, N. E., T. Boyce, M. A. Chesney, S. Cohen, S. Folkman, R. L. Kahn and S. L. Syme (1994), "Socioeconomic status and health", *American Psychologist*, Vol. 49, No. 15, pp. 15-24.

Bloem, S. (2013), "PISA in low and middle income countries", *OECD Education Working Papers*, *93*, OECD Publishing.

Bollen, K. A. (2002), "Latent variables in psychology and the social sciences", *Annual Review of Psychology*, Vol. 53, No. 1, pp. 605-634.

Boyle, M. H. and J. D. Willms (1999), "Place effects for areas defined by administrative boundaries", *American Journal of Epidemiology*, Vol. 149, No. 6, pp. 577-585.

Bourdieu, P. (1977), "Cultural reproduction and social reproduction", in J. Karabel and A. H. Halsey (eds.), *Power and ideology in education*, New York: Oxford University Press.

Bray, M. (2000), "Community Partnerships in Education: Dimensions, Variations and Implications", EFA Thematic Study, Dakar.

Carroll, J. (1963), "A model for school learning", *Teacher College Record*, Vol. 64, pp. 723-733.

Coleman, J. (1988), "Social capital in the generation of human capital", *American Journal of Sociology*, Vol. 94, S95-S120.

Deaton, A. (2002), "Policy implications of the gradient of health and wealth", *Health Affairs*, Vol. 21, No. 2, pp. 13-30.

Edwards, J. R. and R. P. Bagozzi (2000), "On the nature and direction of the relationship between constructs and measures", *Psychological Methods*, Vol. 5, pp. 155–174.

Francis, D. J., et al. (1996), "Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis", *Journal of Educational Psychology*, Vol. 88, pp. 3-17.

Heckman, J. (2008), *Schools, Skills and Synapses*, Bonn, Germany: Institute for the Study of Labor.

Hill, P. and K. Rowe (1996), "Multilevel modelling in school effectiveness research", *School Effectiveness and School Improvement*, Vol. 7, No. 1, pp. 1-34.

Johnson, W. and R. F. Krueger (2005), "Genetic effects on physical health: Lower at higher income levels", *Behavior Genetics*, Vol. 35, No. 5, pp. 579-590.

Klieme, E., C. Pauli and K. Reusser (2009), "The Pythagoras Study: Investigating Effects of Teaching and Learning in Swiss and German Classrooms", in T. Janik and T. Seidel (eds.), *The Power of Video*

Studies in Investigating Teaching and Learning in the Classroom, Waxmann Verlag, Münster, pp. 137-160.

Klieme, E., S. Kuger and members of the PISA 2015 Questionnaire Expert Group (2014), "PISA 2015: Draft questionnaire framework", Frankfurt: German Institute for International Educational Research (DIPF).

Kozminsky, E. and L. Kozminsky (2001), "How do general knowledge and reading strategies ability relate to reading comprehension of high school students at different education levels?", *Journal of Research in Reading*, Vol. 24, No. 2, pp. 187-204.

Lareau, A. M. (1987), "Social class differences in family-school relationships: The importance of cultural capital", *Sociology of Education*, Vol. 60, pp. 73-85.

Lareau, A. M. and E. B. Weininger (2003), "Cultural capital in educational research: A critical assessment", *Theory and Society*, Vol. 32, pp. 567-606.

Levin, H. M. (1989), "Economics of investment in educationally disadvantaged students", *American Economic Review*, Vol. 79, No. 2, pp. 52-56.

Marmot, M. G., G. D. Smith, S. Stansfield, C. Patel, J. Head, I. White, E. Brunner and A. Feeney (1991), "Health inequalities among British civil servants: The Whitelhall II study", *Lancet* Vol. 337, pp. 1387-1393.

Morris, D., J. Bloodgood and J. Perney (2003), "Kindergarten predictors of first and second-grade reading achievement", *The Elementary School Journal*, Vol. 104, No. 2, pp. 93-109.

Mortimore, P., P. Sammons, L. Stoll, D. Lewis and R. Ecob (1988), *School Matters*, Los Angeles, University of California Press.

Mueller, C. W. and T. L. Parcel (1981), "Measures of Socioeconomic Status: Alternatives and recommendations", *Child Development*, Vol. 52, pp. 13-30.

Murillo, F. J. and M. Román (2011), "School infrastructure and resources do matter: analysis of the incidence of school resources on the performance of Latin American students", *School Effectiveness and School Improvement*, Vol. 22, No. 1, pp. 29-50.

National Center for Education Statistics (2012), "Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation", Washington: NCES.

Nonoyama-Tarumi, Y. and J. D. Willms (2010), "The relative and absolute risks of disadvantaged family background and low levels of school resources on student literacy", *Economics of Education Review*, Vol. 29, No. 2, pp. 214-224.

O'Reilly, T. and D. McNamara (2007), "The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional 'high-stakes' measures of high school students' science achievement", *American Educational Research Journal*, Vol. 44, No. 1, pp. 161-196.

OECD (2013a), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, PISA, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264190511-en

OECD (2013b), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices*, PISA, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264201156-en

OECD (2011), *Education at a Glance 2011: OECD Indicators*, OECD Publishing, Paris. http://dx.doi.org/10.1787/eag-2011-en

OECD (2010), *Iberoamerica in PISA 2006: Regional Report*, Santillana Educación, Paris. http://recursostic.educacion.es/inee/pisa/lectora/_private/iberoamericaenpisa2006.pdf

OECD (2001), *Knowledge and Skills for Life: First Results from PISA 2000*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264195905-en

Rose, P. (2003), *Communities, gender and education: Evidence from sub-Saharan Africa: Background paper for 2003 UNESCO Global Monitoring Report*, Centre for International Education, University of Sussex.

Sammons, P., J. Hillman and P. Mortimore (1995), *Key characteristics of effective schools: A review of school effectiveness research*, London, UK: Ofsted.

Scheerens, J., C. Vermeulen and W. J. Pelgrum (1989), "Generalizability of instructional and school effectiveness indicators across nations", *International Journal of Educational Research*, Vol. 13, pp. 789-799.

Scheerens, J. and R. J. Bosker (1997), *The Foundations of Educational Effectiveness*, Oxford, UK: Pergamon.

Slavin, R. E. (1994), "Quality, appropriateness, incentive, and time: a model of instructional effectiveness", *International Journal of Educational Research*, Vol. 21, No. 2, pp. 141-158.

Tramonte, L. and J. D. Willms (2010), "Cultural capital and its effects on education outcomes", *Economics of Education Review*, Vol. 29, No. 2, pp. 200-213.

Willms, J. D. (2014), *Student engagement on the pathway to success: Alberta Education's Tell Them From Me project*, Edmonton: Alberta Education.

Willms, J. D. (2011), *An analysis plan on educational equality and equity: Recommendations for the OECD Education at a Glance*, paper prepared for the OECD NESLI INES Network for the collection and adjudication of system-level descriptive information on educational structures, policies and practices (NESLI), Fredericton: UNB-CRISP.

Willms, J. D. (2010), "School composition and contextual effects on student outcomes", *Teachers College Record*, Vol. 112, No. 4, pp. 1008-1037.

Willms, J. D. (2006), *Learning divides: Ten policy questions about the performance and equity of schools and schooling systems*, Montreal: UNESCO Institute for Statistics.

Willms, J. D. (2003), *Student engagement at school: A sense of belonging and participation*, OECD Publishing.

Willms, J. D., et al (2014), "Assessing educational equality and equity with large-scale assessment data: Brazil as a case study", Washington: Inter-American Development Bank.

Willms, J. D. and L. Paterson (1995), "A multilevel model for community segregation", *Journal of Mathematical Sociology*, Vol. 20, No. 1, pp. 23-40.

Willms, J. D., and M.-A. Somers (2001), "Family, classroom and school effects on children's educational outcomes in Latin America", *International Journal of School Effectiveness and Improvement*, Vol. 12, No. 4, pp. 409-445.

Wolfson, M., et al. (1999), "Relation between income inequality and mortality: Empirical demonstration", *British Medical Journal*, Vol. 319, pp. 953-957.

World Bank (1993), *World Development Report 1993*, New York: Oxford University Press.

Wu, A. D., Z. Li and B. D. Zumbo (2007), "Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS Data", *Practical Assessment Research and Evaluation*, Vol. 12, No. 3, pp. 1-26.

ANNEX A: EVOLUTION OF PISA CONTEXTUAL QUESTIONNAIRES, 2000-2012

Level/Unit	Construct	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA Cycles
Child	**Date of Birth**	S (1)	S (3)	S (3)	S (3)	S (3)	**5**
	Grade	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Gender	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Country Born	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Age when Immigrated		S (1)	S (1)	S (1)	S (1)	**4**
	Attend special courses in school	S (4)	S (2)				**2**
	Attend courses out of school	S (7)	S (3)		S (9) S (6)		**3**
	School Program	S (1)	S (6)	S (6) S (8)	S (6)	S (6)	**5**
	Attend pre-primary		S (1)		S (1)	S (1)	**3**
	Age started Primary		S (1)		S (1)	S (1)	**3**
	Repeated a Grade		S (1)		S (1)	S (1)	**3**
	Attendance	S (3)	S (1)			S (1) S (1) S (1)	**3**
	Sense of Belonging at School	S (8)	S (6)			S (9)	**3**
	Values Schooling Outcomes		S (4)	S (3)	S (4)	S (4) S (4)	**4**
	Learning Efficacy					S (6) S (6) S (6) S (5)	**1**
	Problem Solving Approach					S (4) S (4) S (4)	**1**
	Values Scientific Knowledge			S (10) S (6), S (6) S (5)			**1**
	Regulation of Environment			S (7)			**1**

Level/Unit	Construct	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA Cycles
	Knowledgeable about science			S (5)			1
	Homework Behaviour	S (4), S (3)	S (1)				2
	Engagement in Reading	S (1), S (6)			S (1) S (5) S (7)		2
	Domain-specific Processes for Reading/Math/Science (e.g., motivation, and opportunity to learn)	S (9), S (1)	S (8)	S (5) S (8)	S (11)	S (8) S (6) S (5)	5
	Confidence in Learning Math/Science		S (8)	S (8) S (6)		S (8)	3
	Familiar with Math Terms					S (16)	1
	Anxiety about Math		S (10)			S (10)	2
	Effort in Math					S (9) S (5)	1
	Extra Learning Time – Math/Science		S (6)	S (6) S (6)		S (8) S (4) S (6)	3
	Ways of Learning Math/Reading		S (14)		S (6) S (6)	S (12)	3
	Where learn about science			S (8) S (6)			1
	Preparation and desire for science career			S (4) S (4) S (4)			1
	Study Practices				S (13)		1
	Uses School Resources	S (5)					1
	Use Library				S (7)		1
	Occupational Aspirations	S (1)		S (1)			2
	Educational Aspirations		S (1)				1
	School Marks	S (3)					1
	Why Attend This School?		S (6)				1

Level/Unit	Construct	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA Cycles
Family	**Family Structure**	S (8)	S (5)		S (6)	S (6)	**4**
	Number Siblings	S (1)					**1**
	Mother's Occupation	S (4), S (1), S (1)	S (4), S (1), S (1)	S (2)	S (2) S (1)	S (1) S (1) S (1)	**5**
	Father's Occupation	S (4), S (1), S (1)	S (4), S (1), S (1)	S (2)	S (2) S (1)	S (1) S (1) S (1)	**5**
	Mother's Education	S (1), S (1)	S (5)	S (1) S (3)	S (1) S (4)	S (1) S (4)	**5**
	Father's Education	S (1), S (1)	S (5)	S (1) S (3)	S (1) S (4)	S (1) S (4)	**5**
	Home Possessions	S (7)	S (5)	S (5)	S (5)	S (5)	**5**
	Educational Possessions (and some others)	S (11)	S (16)	S (17)	S (17)	S (17)	**5**
	Books in the Home	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Country born – Mother	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Country born – Father	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Language Spoken at Home	S (1)	S (1)	S (1)	S (1)	S (1)	**5**
	Cultural Communication with Parents	S (3)					**1**
	Social Communication with Parents	S (3)					**1**
	Cultural Capital	S (6)					**1**
	Parental Involvement (general)	S (6)					**1**
	Family Involvement (Schoolwork)	S (6)					**1**
Classroom	**Class Size (3 subjects)**	S (3)	S (1)	S (1)	S (1)	S (1)	**5**
	Number classes (3 subjects)	S (3)	S (3)				**2**
	Learning Time (3 subjects and other)			S (12)	S (3) S (3) S (1)	S (3) S (3) S (1)	**3**

Level/Unit	Construct	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA Cycles
	Achieve Press	S (4)				S (5)	2
	Teacher-student relations	S (5)	S (5)		S (5)	S (5)	4
	Teacher Support	S (5)	S (5)			S (5)	3
	Classroom Disciplinary Climate	S (6)	S (5)		S (5)	S (5) S (3) S (4)	4
	Homework Behaviour	S (3)				S (3)	2
	Competitive Learning - Math		S (5)				1
	Cooperative Learning - Math		S (5)				1
	Math exposure to problems					S (9) S (8)	1
	Instructional Approach			S (17)	S (7) S (9)	S (13) S (9)	3
School	**Type of Community (rural/urban)**	A (1)	A (1)	A (1)	A (1)	A (1)	5
	Enrollment	A(2)	A(2)	A(2)	A(2)	A(2)	5
	Class Size (Test Language)					A (1)	1
	School Type	A (1)	A (1)	A (1)	A (1)	A (1)	5
	Grades covered	A (14)	A (14)	A (14)	A (14)		4
	Grade Repetition		A (2)	A (2)	A (2)		3
	Dropout Rate					A (1)	1
	Funding Sources	A (4)	A (4)	A (4)	A (4)	A (4)	5
	Administrator's Sex				A (1)		1
	Instructional Time (by program)	A (3)	A (3)				2
	Selective School	A (7)	A (7)	A (6)	A (6)	A (7)	5
	School Transfer out	A (6)			A (6)	A (6)	3

Level/Unit	Construct	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA Cycles
	School Programs	A (6)					**1**
	Programs for 2nd Language Students				A (5)		**1**
	Program Choice	A (5)					**1**
	Infrastructure	A (9)	A (20)	A (13)	S (1) A (13) A (14)	A (13) A (11)	**5**
	Staff Resources	A (5)					**1**
	Computing Resources	A (6)	A (6)	A (3)	A (3)	A (3)	**5**
	Expectations for Computer Use					A (3)	**1**
	Teacher FTE	A (9)	A (6)	A (6)	A (6)	A (6)	**5**
	Math/Science Teacher FTE		A (10)	A (1)		A (10)	**3**
	Teacher PD	A (1)				A (2)	**2**
	Math Teacher Monitoring		A (4)			A (4)	**2**
	Innovative Practices (Math)		A (3)				**1**
	Standardized Practices (Math)					A (3)	**1**
	Extra Math Programs					A (1) A (1)	**1**
	Science Extra-curricular			A (5) A (1) A (1) A (1)			**1**
	Science offerings on Environment			A (4) A (5)			**1**
	Teachers encourage science careers			A (1) A (1) A (1)			**1**
	Career Guidance (general)			A (1)			**1**
	Assessment Frequency	A (5)	A (5)		A (5) A (3)		**3**

Level/Unit	Construct	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA Cycles
	Assessment Reporting Frequency	A (3)					1
	Assessment Reporting to Parents			A (3)	A (3)		2
	Use of Assessment	A (6)	A (8)	A (5)	A (8) A (5)	A (8) A (2)	5
	Parental Expectations			A (3)	A (3)	A (3)	3
	Parental Choice of Schools			A (3)	A (3)	A (3)	3
	Parental Involvement					A (12)	1
	Barriers to Student Learning	A (17), A (5)	A (13)		A (13)	A (19)	4
	Students' first language		A (4) A (2)		A (1)		2
	Streaming (3 subjects)		A (4)	A (2)	A (3)	A (4)	4
	Engagement (math)		A (5)				1
	Teachers' Attitudes	A (4)	A (4) A (3) A (3)			A (4)	3
	Teachers' Attitudes (Math)					A (6)	1
	Teacher Accountability					A (7)	1
	Students' Attitudes		A (7)				1
	Principals' Management				A (14)	A (18) A (3)	2
	Quality Assurance					A (9)	1
	Autonomy	A (12)	A (12) A (7)	A (12) A (6)	A (12) A (6)	A (12)	5

ANNEX B: CONSTRUCTS MEASURED IN OTHER INTERNATIONAL SURVEYS

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
Child	Date of Birth	5							
	Grade	5							
	Gender	5							
	Country Born	5							
	Age when Immigrated	4							
LHS	Other languages			1		1			
LHS	Familiarity with language of test		1						
SES	Child Labor				2	5			18
SES	Hazardous labor conditions								6
SES, LT	Hours of labor per week					1			2
SES	Nutrition						3		
SES	Health/wellbeing			13	6		1		
SES	Time and means to reach school				1				

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
SES	Place to stay during school week						1		
	Attend special courses in school	2							
	Attend courses out of school	3							
	School Program	5							
	Attend pre-primary	3							
	Age started Primary	3							
ELO	School readiness		11						
	Repeated a Grade	3							
	Attendance	3							
LT	Dropout				4				
	Sense of Belonging at School	3							
	Values Schooling Outcomes	4							
	Learning Efficacy	1							
	Problem Solving Approach	1							

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
LT	Values Reading		6						
	Values Scientific Knowledge	1							
	Regulation of Environment	1							
	Knowledgeable about science	1							
	Homework Behaviour	2							
	Engagement in Reading	2							
LT	Enjoys reading		7						
	Attitude towards Reading/Math/Science	5							
	Confidence in Learning Math/Science	3							
	Familiar with Math Terms	1							
	Anxiety about Math	2							
	Effort in Math	1							
	Extra Learning Time – Math/Science	3							
	Ways of Learning Math/Reading	3							

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
	Where learn about science	1							
	Preparation and desire for science career	1							
	Study Practices	1							
	Uses School Resources	1							
	Use Library	1							
	Occupational Aspirations	2							
	Educational Aspirations	1							
	School Marks	1							
	Why Attend This School?	1							
RES	Access and use of technology in and outside school		1	19					
Family	Family Structure	4							
	Number Siblings	1							
	Mother's Occupation	5							

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
	Father's Occupation	5							
	Mother's Education	5							
	Father's Education	5							
SES	Literate parents				2	2			2
	Home Possessions	5							
	Educational Possessions (and some others)	5							
SES	Basic educational resources					6	9		
	Books in the Home	5							
	Country born – Mother	5							
	Country born – Father	5							
LHS	**Language Spoken at Home**	5							
SES	Parental reading practices		9			1			1
SES	Parental access to media and computer								8
ELO	Reading to the child		15			1			1

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
	Cultural Communication with Parents	1							
	Social Communication with Parents	1							
FAM	Social Capital			13					
	Cultural Capital	1							
	Parental Involvement (general)	1							
	Family Involvement (Schoolwork)	1							
FAM	Parental disciplinary practices								13
SES	Household composition			4	3	4			5
SES	Household income/money transfer			9	3				
SES	Dwelling infrastructural features				6	11			4
SES	Property of the dwelling				1				1
SES	Property of farming land or animal				4				4
SES	Water and sanitation				20				2
SES	Electricity				2				2

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
Classroom	**Class Size (3 subjects)**	5							
	Number classes (3 subjects)	2							
	Learning Time (3 subjects and other)	3							
RES	Use of learning material		16			5		X	
	Achieve Press	2							
	Teacher-student relations	4							
	Teacher Support	3							
	Classroom Disciplinary Climate	4							
	Homework Behaviour	2							
QI, RES	Learning basic practices					10			
QI, RES	Reading: teaching activities		20				8		
QI, RES	Math: teaching activities								
QI, RES	Reading: teaching approach						5		
QI, RES	Reading: teacher's goal setting						7		

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
	Competitive Learning - Math	1							
QI	Cooperative Learning - Math	1							
QI	Math exposure to problems	1							
	Instructional Approach	3							
LT	Curriculum coverage		26					X	
RES	Math assessment							X	
RES	Reading assessment		8					X	
RES	Differentiation/scaffolding							X	
RES	Active learning							X	
School	**Type of Community (rural/urban)**	5							
	Enrollment	5							
RES	Percentage of poor children		1					X	
RES	Percentage languages different from that of the test		1						

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
RES	Transportation/geographical accessibility							X	
	Admission policies							X	
	Class Size (Test Language)	1							
	School Type	5							
	Grades covered	4							
	Grade Repetition	3							
	Dropout Rate	1							
	Funding Sources	5							
	Administrator's Sex	1							
RES	Teacher/Admin demographics		2				4	X	
RES	Teachers and Admin living conditions						7		
RES	Teachers and Admin home possessions						14		
RES	Teacher and Admin satisfaction with living conditions						6		

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
RES	Teacher education		2				12	X	
RES	Teacher years of experience		1				3	X	
	Instructional Time (by program)	2							
	Selective School	5							
	School Transfer out	3							
	School Programs	1							
	Programs for 2nd Language Students	1							
	Program Choice	1							
	Infrastructure	5							
RES	Safety		2				2		
	Bullying		6					X	
	Staff Resources	1							
	Violence		9			6	1	X	
	Students' Absenteeism		1					X	

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
RES	Teachers' negative behaviours		2				2		
RES	Teachers' job satisfaction		7				11	X	
RES	Teacher/Admin minimal-basic resources		6				28	X	
RES	Labs/Area-specific resources		2					X	
	Computing Resources	**5**							
	Expectations for Computer Use	1							
RES	Remedial activities							X	
	Teacher FTE	**5**							
	Math/Science Teacher FTE	3							
	Teacher PD	2							
	Math Teacher Monitoring	2							
RES	Teachers collaboration/support		5					X	
	Innovative Practices (Math)	1							
	Standardised Practices (Math)	1							

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
	Extra Math Programs	1							
	Science Extra-curricular	1							
	Science offerings on Environment	1							
	Teachers encourage science careers	1							
RES	Assessment of teachers' practices		8						
	Career Guidance (general)	1							
	Assessment Frequency	3							
	Assessment Reporting Frequency	1							
	Assessment Reporting to Parents	2							
	Use of Assessment	5							
FAM	Communication and involvement of family		9					X	
	Parental Expectations	3							
	Parental Choice of Schools	3							
	Parental Involvement	1							

Level/Unit	Construct	PISA Cycles	PIRLS (IEA)	LAMP (UNESCO)	STEP (World Bank)	LLECE (PERCE SERCE) TERCE	PASEC and SACMEQ	WEI	MICS
	Barriers to Student Learning	4							
	Students' first language	2							
	Streaming (3 subjects)	4							
	Engagement (math)	1							
	Teachers' Attitudes	3							
	Teachers' Attitudes (Math)	1							
	Teacher Accountability	1							
	Students' Attitudes	1							
RES	Economic incentives for teachers		1						
RES	Leadership activities		13					X	
	Principals' Management	2							
	Quality Assurance	1							
	Autonomy	5							
	Community activity							X	

ANNEX C: LIST OF INTERNATIONAL STUDIES REVIEWED

PIRLS: Progress in International Reading Literacy Study.

PIRLS is an international comparative study of the reading literacy of 4th-grade (or equivalent) students in participating countries and is co-ordinated by the International Association for the Evaluation of Educational Achievement (IEA).

The first administration of PIRLS was in 2001, and included 36 education systems (including countries and subnational entities, such as Canadian provinces and Hong Kong, a Special Administrative Region of the People's Republic of China). The second administration in 2006 included students from 45 education systems. The third administration of PIRLS was in 2011, with 53 education systems participating.

LAMP: Literacy Assessment and Monitoring Programme

LAMP measures reading and numeracy skills among youth and adults living in countries at different stages of economic development. The programme conducts assessment and data collection in ten different languages, to produce indicators that can be compared across time, countries and cultures. The UIS has developed LAMP to improve literacy skills by monitoring the distribution of reading and numeracy skills within the youth and adult populations. LAMP allows communicating to the countries the information to plan and carry out initiatives to improve literacy skills; it also developed standardised methods and tools to measure reading and numeracy skills and build analytical skills and competencies in the participating countries.

STEP: The Skills Toward Employment and Productivity Skills Measurement Study

In 2010, World Bank launched STEP Skills Measurement Study in developing countries to identify the current levels and distributions of cognitive, technical and non-cognitive skills among adults in middle and low-income countries. STEP aims at assessing the mismatch between the skills of the adult population and the needs of employers as well as the impact of skills on labour market outcomes.

STEP is meant to provide guidance to design and implement skills-enhancing interventions to improve countries' employment and productivity.

LLECE: Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación

The LLECE included three cycles:

- Primer Estudio Regional Comparativo y Explicativo (PERCE) (1997)

- Segundo Estudio Regional Comparativo y Explicativo (SERCE) (2006)

- Tercer Estudio Regional Comparativo y Explicativo (TERCE) (2013)

LLCE is the network that evaluates and monitors the quality of education in Latin America. It is co-ordinated by the Oficina Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC) in Santiago de Chile. LLECE produces information and knowledge for policy design and

implementation in education. PERCE, SERCE, and TERCE studies are part of the LLCE/OREALC action plan. The three assessments involved students in grade 3 and grade 6 and provided data and knowledge on how to offer students universal and equal access to pertinent education. PERCE, SERCE, and TERCE inform policy makers, school administrators, and teachers on equality, equity, and potential for intervention.

PASEC: CONFEMEN Programme for the Analysis of Education Systems

The CONFEMEN has implemented PASEC since 1991. The analytical framework includes the definitions of literacy and mathematics, the domains in literacy (6 in 2nd grade; 2 in 5th grade), and the domains in mathematics (2 in 2nd grade; 4 in 5th grade). Since 2012 PASEC has started to focus on 6th grade students for the upper level in order to make assessment results comparable with SACMEQ. For the contextual data collection PASEC has questionnaires for students, teachers and principals. Since PASEC was implemented in Senegal from 2005 to 2006 it would be important to note any country specific adaptations of this assessment like adaptations based on the national curriculum and/or translation into the local languages.

SACMEQ: Southern and Eastern Africa Consortium for Monitoring Educational Quality

The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) is an international non-profit developmental organisation of 15 Ministries of Education in Southern and Eastern Africa committed to support education planners in learning how to apply scientific methods to monitor and evaluate the conditions of schooling and the quality of education. The SACMEQ I, II, and III focussed on an examination of the conditions of schooling in relation to achievement levels of learners and their teachers in reading, and mathematics. The assessment includes also the learners' levels of basic knowledge about HIV and AIDS.

WEI: The World Education Indicators

The WEI programme was founded in 1997 as a joint endeavour of the UNESCO Institute for Statistics (UIS) and the Organisation for Economic Co-operation and Development (OECD). It aims to consolidate basic education statistics and track the recurrence or emergence of new education issues.

WEI-SPS: The WEI Survey of Primary Schools

This study was established in 2002 to address issues related to quality and equity of education by producing internationally comparable data on school processes and infrastructure, on teachers and teaching, and on learning conditions and resources. WEI-SPS collected detailed information on the context, conditions and conduct of primary schools. For some participating countries, the study also presented an opportunity for capacity building to support large-scale national assessments.

MICS: Multiple Indicators Cluster Survey

UNICEF assists countries in collecting and analysing data to fill data gaps for monitoring the situation of children and women through its international household survey initiative the Multiple Indicator Cluster Surveys (MICS). MICS harmonises a wide range of indicators in the areas of health, education, child protection, water and sanitation and HIV and AIDS. At the country level, MICS findings are a basis for policy decisions and programme interventions, and for influencing public opinion on the situation of children and women. MICS surveys are usually carried out by government organisations, with the support and technical assistance of UNICEF.