6

PISA 2015 context questionnaires framework

This chapter describes the core content of the Programme for International Student Assessment (PISA) 2015 and PISA's interest in measuring student's engagement at school, dispositions towards school and their self-beliefs, and in gathering information about students' backgrounds and the learning environment at school. The chapter discusses the content and aims of the Student Questionnaire, the School Questionnaire (completed by school principals), the optional Parent Questionnaire (completed by parents of students who sat the PISA test), the optional Educational Career Questionnaire (completed by students, concerning their educational and career aspirations), the optional ICT Familiarity Questionnaire (completed by students, concerning their attitudes towards and experience with computers) and the optional Teacher Questionnaire (completed by teachers, and introduced in PISA 2015).

Providing indicators on the effectiveness, equity and efficiency of education systems, setting benchmarks for international comparison, and monitoring trends over time are the most important goals of the Programme for International Student Assessment (PISA). In addition, PISA builds a sustainable database that allows researchers worldwide to study basic and policy-oriented questions on education, including its relation to society and the economy.

In order to reach these goals, PISA not only needs reliable and valid measures for assessing cognitive student achievement (in reading, mathematics, science literacy and other "life skills"), but also information on non-cognitive outcomes (e.g. students' motivation to learn), individual conditions (e.g. students' cultural, ethnic and socio-economic background), and the structural and process characteristics of the institutional context (e.g. teaching practices and learning opportunities in classrooms, leadership and school policies for professional development, vertical and horizontal differentiation of the school system). It is this diverse set of constructs that is measured through questioning various stakeholders, namely students and school principals, and (on an optional basis) parents and teachers.

PISA 2015 is the sixth wave of the programme. Since 2000, the so-called "Background Questionnaires" have gained substantially in importance. Meanwhile, questionnaires are of interest in their own right, beyond providing the "background" for reporting test results. PISA 2015 is a combination of student assessment in selected areas of achievement and context assessment on different policy-related issues. One reason for this change of perspective is that policy makers wanted the programme to report on a variety of issues that are relevant to professional practice, governance and policy making in education. Thus, the scope of issues to be covered has gradually expanded since PISA 2000. In addition to student achievement and its relation to individual background, learning opportunities and non-cognitive outcomes, educational policies and practices are monitored through indicators. Also, data analysis and reporting have become more complex, allowing for more in-depth reporting. In addition to providing tables of indicators, patterns of input, process and outcome variables are identified within and across countries; trends are reported, relations are explored, impact is estimated. These analyses require more sophisticated modelling approaches, and detailed data on contextual factors regarding students, schools and education systems.

Probably the most important characteristic of PISA, after 15 years of existence, is the availability of system-level trend data. PISA allows for the description of change in a country's performance level over time, but also for the description of changes in non-cognitive outcomes, living conditions of adolescents and their families, professional practices and organisational structures for schooling. The more PISA moves towards repeated cycles of measurement, the more can be learned from examining the stability and variability of conditions, processes, outcomes and their relations: policy makers can use trend data for constant diagnosis and feedback; the explanatory power of the study will increase because changes in performance can be interpreted and explained more substantively, taking changes in input and process into account (Gustafsson, 2008; Hanushek and Wößmann, 2011); and trend analyses are less likely to be biased by cultural issues.

Quite often, policy makers and researchers have been reluctant to interpret "soft" constructs, such as school climate, students' motivation, teachers' job satisfaction or parents' self-reported engagement, fearing that they are not sufficiently comparable across countries. Now that trends are available, the focus is on change rates within countries rather than on cross-sectional comparisons of status. For example, the question of whether well-being in schools is increasing or decreasing will be a relevant indicator within countries, and this indicator is not affected by response styles differing between countries. However, trend analyses require PISA to define a general set of constructs that will remain constant over several cycles in the future. This set of constructs will be referred to as "core content" in this chapter.[1]

This framework intends to explain the goals and the rationale for choosing appropriate questionnaire content, guiding both questionnaire development and the forthcoming reports. The document is organised into two main parts: a definition of the core content of the PISA questionnaires, and a description of the modular structure for broader coverage of policy issues.

The first part of this document links the current framework to the overarching (cross-cycle) structure of PISA context assessment set out in the PISA 2012 Framework (OECD 2013: 168 ff.). The constructs that need to be covered for monitoring trends in education are revisited, with reference to the general background of Educational Effectiveness Research. Measures that have been used previously for initial reporting, for international indicators and for secondary analyses are overviewed, culminating in an outline of core content that should be assessed in all cycles, for all participants.

The second and larger part of this document explores the full breadth of policy issues to be covered, structured in 19 modules, and explains how the most important modules – i.e. those judged as high priority modules by the PISA Governing Board – have been implemented in PISA 2015. Here, detailed references to current research are provided.

The development of questionnaire content for PISA 2015 has been challenged not only by the intention of covering a broad array of policy issues, but also by the introduction of Computer-Based Assessment (CBA) as the preferred mode of administration for both cognitive tests and questionnaires. In addition to developing new questions, all content from previous cycles had to be revisited, sometimes transformed into a CBA-friendly format (like using a "slider" that can be manipulated interactively instead of having the test-taker fill in numeric information), and field trialled anew. Finally, the analysis of process data based on logfile recording will deepen the understanding of response behaviour.

Based on careful analysis of the field trial data and thoughtful discussion of priorities among experts and policy makers, constructs, questions and items were selected for inclusion in the PISA 2015 main survey. The Student Questionnaire was targeted to last 35 minutes, on average.

DEFINING THE CORE OF CONTEXT ASSESSMENT IN PISA

Choosing from the many measures that might be incorporated in the PISA study design is a complex process, directed by the priorities that countries have set for the study, but also informed by education research. One of the major forces that drive the PISA design in general is the rhythmic change of focus in the cognitive assessment: reading literacy has been or will be the major domain of assessment in PISA 2000, 2009 and 2018; mathematics is the focus of PISA 2003, 2012 and 2021; science takes the lead in PISA 2006, 2015 and 2024. Whatever serves as the major domain of cognitive assessment shall also be a major focus of "domain-specific" context assessment. However, there is a need for some stability in measurement to understand trends in education.

The Questionnaire Framework for PISA 2012 established an overarching structure that delineated core questionnaire content that should be kept comparable across cycles (OECD 2013: 189 ff.) to allow for continuous monitoring of education systems. The overarching framework refers to domain-specific as well as domain-general measures assessing conditions, processes and outcomes of education for both individual students and schools. Finding an appropriate balance among these facets of the design is crucial for the long-term success of the PISA programme. In order to establish valid and reliable trends at the country level, it is important to implement a stable set of variables, which will be used as major reporting variables across PISA cycles.[2]

This overarching framework is discussed below, along with specifying the constructs and measures in more detail and providing arguments that support the choice of core content for PISA 2015.

Outline of core content: Constructs to be covered

Taking into account the goals of context assessment in PISA as stated in the introduction, strategic decisions made by the PISA Governing Board, the overarching framework developed for PISA 2012, and recommendations from the research literature, the current framework assumes that education policy makers in participating countries need to be informed in four broad areas: outcomes, student background, teaching and learning processes, and school policies and educational governance. These areas are described below. As stated above, the following sections elaborate in more detail what has already been established in the Questionnaire Framework for PISA 2012.

Non-cognitive outcomes

The main challenge of PISA concerns measuring and documenting the outcomes of education that have been reached up to the age of 15 years. Educating a person basically means fostering his or her individual development as a unique, self-determined, knowledgeable person who gradually gains in ability to participate in society. As each PISA assessment is a cross-sectional study, PISA cannot capture developmental processes; but PISA serves as a snapshot of development status at the age of 15.

This snapshot includes an assessment of literacy and life skills, but in addition to these cognitive outcomes, other factors are important, too. Success in school – and in life – depends on being committed, sharing values and beliefs, respecting and understanding others, being motivated to learn and to collaborate, and being able to regulate one's own learning behaviour. These constructs can be perceived as prerequisites of cognitive learning, but may also themselves be judged as goals of education, as the OECD project *Defining and Selecting Key Competencies* (DeSeCo) has elaborated (Rychen and Salganik, 2003). Education research and econometric analyses have shown that non-cognitive factors are most important for individual development as well as for success in life and well-being, and thus have an impact on individuals and society alike (Heckman, Stixrud and Urzua, 2006; Almlund et al., 2011).

Therefore, PISA addresses non-cognitive outcomes like attitudes, beliefs, motivation and aspirations, and learning-related behaviour, such as invested learning time. These non-cognitive outcomes are measured within the Student Questionnaire (StQ), but also in the ICT Familiarity Questionnaire (ICTQ). They may be of a general nature, such as achievement

motivation and well-being of students, or related to the domains of the cognitive assessment, such as reading engagement, interest in mathematics or enjoyment of science. Domain-specific non-cognitive outcomes are also mentioned in the respective definitions of literacy, so this array of constructs serves as a link between test frameworks and context framework. Students' self-efficacy beliefs – i.e. the strength of their belief in being able to solve tasks similar to the ones tested in the cognitive PISA tests – have been shown to be a strong correlate of student achievement both within and between countries.

Student background

In order to understand education careers and to study equity issues within and across countries, family background variables, such as socio-economic status and immigrant background, have to be taken into account. The distribution of education opportunities and outcomes depending on these background variables shows whether countries succeed in providing equal opportunities.

PISA has become famous for its detailed, theory-based assessment of family background, socio-economic status and immigration background. A lot of effort went into the definition and operationalisation of individual student background indicators, finally leading to the establishment of a powerful, integrated indicator for students' economic, social and cultural status (ESCS; Willms, 2006). The components of this indicator need to be assessed in as stable a way as possible across the PISA cycles. In addition, information on parental support helps understanding of how formal education and family background interact in promoting student learning.

Furthermore, PISA gathers retrospective and prospective information about education pathways and careers across the lifespan. In recent years, researchers and others have stressed the importance of early childhood education (Blau and Curie, 2006; Cunha et al., 2006). Therefore, PISA intends to capture at least some information on primary and pre-primary education.

On top of individual student background, the social, ethnic and academic composition of the school he or she attends has an impact on learning processes and outcomes. Therefore, PISA uses aggregated student data to characterise background factors on the school level in addition to structural factors, such as school location, school type and school size.

Teaching and learning

School-based instruction is the core process of formal, systematic education. Therefore, policy makers need information on teaching, learning and the organisation of schools. To increase the explanatory power of the study, assessment of teaching and learning focuses on the major domain of assessment, which, in 2015, is science. The knowledge base of educational effectiveness research (Scheerens and Bosker, 1997; Creemers and Kyriakides, 2008) allows for the identification of core factors: teachers' qualifications, teaching practices and classroom climate, learning time and learning opportunities provided both within and out of school. For teaching processes, the focus should be on three basic dimensions (Klieme, Pauli and Reusser, 2009): structure and classroom management; teacher support; and cognitive challenge. Addressing teacher- and teaching-related factors in PISA is a challenge, because sampling is by age rather than by grade or class. Nevertheless, aggregated student data and School Questionnaires can serve to describe the learning environment offered by schools.

School policies and governance

As policy makers have limited direct impact on teaching and learning processes, information on school-level factors that help to improve schools, and thus indirectly improve student learning, have high priority. As with teacher and teaching variables, school-effectiveness research has built a strong knowledge base showing that "essential supports" promote school effectiveness (Bryk et al., 2010; see also Creemers and Reezigt, 1997; Scheerens and Bosker, 1997): professional capacity, with a focus on professional development; a well-organised curriculum; leadership and school management; parental involvement; school climate (truthful interactions between stakeholders, clear norms and shared values, high achievement expectations); and the use of assessment and evaluation for improvement. These factors are addressed in the PISA questionnaires as domain-general processes on the school level. Also covered is school-level support for teaching the major domain, such as the provision of laboratory space, information and communication technology (ICT), and a school curriculum for science education.

To meet policy requests directly, PISA also needs to address issues related to governance on the system level (Hanushek and Wößmann 2011; Wößmann et al., 2007). "Locus of decision making" measures and accountability practices describe main aspects of governance, namely the distribution of power and control between central and local stakeholders. Allocation, selection, and assessment and evaluation are the basic processes that policy makers and/or school administrators use to control school quality, and to monitor and foster school improvement. Some of this information can be gained from other sources (as documented in *Education at a Glance*), some can be assessed through the PISA School Questionnaire.

Previous use of PISA context data: Measures that were important for analysis and reporting

In order to evaluate the importance of questionnaire content for PISA, it is worthwhile to look at previous cycles and how their data fed into analysis and reporting. Thus, the relevance of specific measures for policy making and research can be taken into account, in addition to the more abstract constructs mentioned before.

PISA questionnaire data have been used for several types of analyses and reports in addition to OECD reports, for instance to construct education indicators (e.g. *Education at a Glance*) and in scientific research papers. Box 6.1 presents the questionnaire material used in the PISA 2009 report, including non-cognitive outcomes and the impact of background variables, individual and school characteristics, processes and policies as well as system-level factors.

Box 6.1 **Measures based on questionnaires used in *PISA 2009 Results: What Students Know and Can Do***

Volume I: *Student Performance in Reading, Mathematics and Science*

Student background: Gender

Volume II: *Overcoming Social Background: Equity in Learning Opportunities and Outcome*

Student background: ESCS, gender, immigration status, language spoken at home, age of arrival, country of origin

Individual support assessed through Parent Questionnaire: Parental support (at beginning of primary education/ at age 15), pre-primary education (attendance, quality)

Volume III: *Learning to Learn: Student Engagement, Strategies and Practice*

Student background: ESCS, gender, immigration status, language spoken at home

Outcomes: Enjoyment of reading, time and material used for reading, metacognition (awareness of strategies), self-reported use of reading strategies (memorisation, elaboration, control)

Volume IV: *What Makes a School Successful? Resources, Policies and Practices*

Student background: socio-economic status, age of school entry, grade repetition

Student-reported processes: learning time (previous education, learning time at school, enrichment/remedial education, after-school lessons), teacher-student relations, disciplinary climate, teacher's stimulation of reading engagement

School input, policies and processes (reported by the principal): type of school (public/private), number of programmes, class size, educational resources (e.g. ICT, library), school responsibility for assessment and curriculum/for resource allocation, extracurricular activities provided, school admittance/grouping/transfer policies, assessment practices/purposes, use of achievement data, school accountability, methods for monitoring teachers, teacher and student behaviour, parent involvement and expectations, leadership.

Source: OECD, 2010a, 2010b, 2010d and 2010e.

The initial report on PISA 2006 studied the impact of schools on student outcomes, while many papers listed on the Education Resources Information Center (ERIC) international data base discuss non-cognitive, domain-specific outcomes using PISA 2006 data. Seventeen of the publications that used multivariate analyses are listed in Appendix 1.

Selecting and organising the core content

Addressing policy needs, and covering measures that have been used for reporting in previous cycles, a selection of core questionnaire content for PISA 2015 and beyond can be proposed. Figure 6.1 organises the suggested content according to the model that has informed the design of international large-scale assessments for a long time (see e.g. Purves, 1987; OECD 2013: 173 ff.). The model allocates background, process and outcome characteristics of education at respective levels of action (i.e. the system level, the school level, including instruction/class/teacher factors, and the individual student level).

Figure 6.1 ■ **Measures to be included in the core context assessment for PISA**

	Student and school background	Processes	Non-cognitive outcomes
System level		**Governance:** Decision making, horizontal and vertical differentiation	(aggregated student data)
School level	School location, type and size of school, amount and source of resources (incl. ICT) Social/ethnic/academic composition Class size, teacher qualification	**School policies:** Programmes offered, admission and grouping policies, allocated learning time, additional learning time and study support, *extracurricular activities*, professional development, leadership, parental involvement, assessment/evaluation/accountability policies, school climate (teacher and student behaviour) **Teaching and learning:** Disciplinary climate, teacher support, *cognitive challenge*	(aggregated student data)
Student level	Gender, socio-economic status (parents' education and occupation, home possessions, number of books at home), language and migration background, grade level, pre-primary education, age at school entry	Grade repetition, programme attended, learning time at school (mandatory lessons and additional instruction), *out-of-school learning*	Domain-general non-cognitive outcomes (e.g. achievement motivation, well-being in school) Domain-specific non-cognitive outcomes (*motivation, domain-related beliefs and strategies, self-related beliefs, domain-related behaviour*)

Note: Measures in italics are adapted to the respective major domain, e.g. science in PISA 2015.

The set of measures included in Figure 6.1 comprises a core context design that covers all construct areas mentioned above, i.e. non-cognitive outcomes, student background, teaching and learning, school policies and governance; and allows for reporting all the analyses that have been included in initial reports, conducting all the research mentioned in Appendix 1, and calculating all indicators that have been developed for *Education at a Glance* (see above)[3]. Figure 6.1 includes all questionnaire indices that have been shown to be strongly correlated with PISA achievement measures (e.g. number of books at home, socio-economic status, self-efficacy, and disciplinary climate), and thus will be instrumental in estimating test scores in PISA ("plausible values"). Therefore, this set of measures is considered for use in further PISA cycles including and beyond PISA 2015. Keeping this core design stable across cycles will enable trend analyses and complex modelling of system-level changes.

Most of the measures mentioned in Figure 6.1 have already been used in previous cycles, mainly in PISA 2006 or PISA 2012, and thus they represent "trend" content that may be kept constant in the future. This includes the science-specific measures from PISA 2006. When reading and mathematics were the major domain of assessment (PISA 2009 and 2012, respectively), different measures were used to represent the same overarching constructs:

- **Cognitive challenge** in classrooms has been represented by teachers' stimulation of reading engagement (2009), opportunities-to-learn (OTL) question types and experience with applied mathematical tasks (2012), and inquiry-based teaching and learning (2006, 2015).
- **Student motivation** has been operationalised by enjoyment of reading (2009), interest in mathematics (2012), and enjoyment of science (2006, 2015).
- **Domain-related behaviour** has been represented by reading for school and diversity of reading material (2009), mathematics work ethics and math behaviour (2012), and media-related science activities (2006, 2015).
- **Domain-related beliefs and strategies** have been represented by subjective norms on mathematics (2012) and environmental awareness and optimism (2006, 2015); self-related beliefs have been represented by mathematics self-efficacy (2012) and science self-efficacy (2015). PISA 2009 introduced a measure of metacognition instead of reading-related beliefs.

EXPANDING THE FRAMEWORK FOR BROADER COVERAGE OF POLICY ISSUES

Modular approach to the PISA design

When the contractor for questionnaire development in PISA 2015 and the Questionnaire Expert Group started their work, they revisited the content areas described above – non-cognitive outcomes, student background, teaching and learning, school policies and governance – and further differentiated them into 19 more fine-graded "modules", which were approved by the PISA Governing Board (PGB) at its meeting in October 2011 – as the building blocks of the PISA 2015 design for context assessment. Figure 6.2 provides a schematic overview of this modular structure, positioning the modules within the overarching structure of background, process and outcome characteristics.

Figure 6.2 ▪ **Modular structure of the PISA 2015 context assessment design**

	Student background		Processes			Non-cognitive outcomes
	Family	Education	Actors	Core processes	Resource allocation	
Science-related topics		5. Out-of-school science experience	1. Teacher qualification and professional knowledge	2. Science teaching practices	12. Learning time and curriculum	4. Science-related outcomes: motivation, interest, beliefs…
				Teaching and learning		
				3. School-level learning environment for science		
General topics	7. Student SES and family 8. Ethnicity and immigration	9. Educational pathways in early childhood	14. Parental involvement 15. Leardership and school management	13. School climate: interpersonal relations, trust, expectations	16. Resources	6. Career aspirations 10. General behaviour and attitudes 11. Dispositions for collaborative problem solving
				School policies		
			17. Locus of decision making within the school system	19. Assessment evaluation and accountability	18. Allocation, selection and choice	
				Governance		

Columns one and two summarise student background characteristics related to their family and the education they received, the three columns in the middle refer to education processes at different levels (system governance, school policies, teaching and learning), and the columns on the right list various outcomes of education. In Figure 6.2, the lower part deals with domain-general topics, while the upper part includes modules that mainly deal with domain-specific (in this case: science-related) topics, including the learning environment at the school level that specifically supports science education (Module 3), such as laboratories, science-related school curricula, collaboration among science staff, and the value attributed to science within the school community. Thus, Figure 6.2 illustrates the combination of domain-general and domain-specific approaches to international large-scale assessment that is typical for all PISA cycles, with either science, reading or mathematics as the major focus of assessment. As PISA integrates cross-curricular achievement measures, like problem solving (in 2012) or collaborative problem solving (in 2015), appropriate non-cognitive outcomes are added (Module 11).

Traditionally, PISA treats the standard questionnaires (School Questionnaire and Student Questionnaire) separately from optional questionnaires that countries may choose to implement or not. PISA 2015 also keeps these questionnaires separate from an operational and reporting point of view, but the Questionnaire Expert Group intended to make the connections between standard and optional questionnaires as transparent as possible. All modules are included in the standard questionnaires to some extent, while the optional questionnaires are used to treat some modules in depth: Educational Career questions address Modules 2, 9, 12 and 14, while the ICT Familiarity Questionnaire contributes to Modules 7, 10 and 16, and the Parent Questionnaire provides content for Modules 5, 8, 9 and 14. The Teacher Questionnaire, which has been added to the design for PISA 2015, is relevant for Modules 1, 2, 11, 12, 15, 16 and 19. Thus, countries opting for any of these additional questionnaires will have additional information available for in-depth analysis of respective policy issues.

The expanded model will guide analysis and reporting in a systematic way:

- Each of the modules can be conceived as a thematic focus for analysis, as will be shown below. Based on a comprehensive review of the corresponding research literature, each module covers main components that are relevant for a specific field of education practice or policy making. Information gathered from students, school leaders and – for countries choosing those options – parents and teachers can be combined to understand patterns and relations within countries and to compare between systems.
- Equity issues in education can be researched by studying outcomes in relation to background factors. (Un-)equal opportunities can be researched by studying schooling offered to various subgroups of students, while efficiency can be described as the relation between outcomes and resources.
- Models of educational effectiveness can be specified and tested by linking schooling to education outcomes, controlling for background factors.

Every module represents a focus of policy making. Thus, the set of 19 modules covers a wide array of policy issues that are relevant across countries. This set is comprehensive, as can be seen by comparing the modular structure with literature on education policy. For example, most topics treated by Sykes, Schneider and Plank (2009) in their state-of-the-art review of education policy research are covered here.

To sum up, the modular approach to context assessment in PISA 2015 allows for a broad coverage of policy issues and related research questions. However, the PISA design sets strict limits on the length of questionnaires; to cover the requested breath of concepts, only some modules or constructs within modules can be focused in more detail. To find relevant points of interest, PGB members were asked to indicate the top priority modules for further development work, based on policy relevance and the need for improvement from previous cycles. More emphasis has been devoted to those areas identified as priorities.

The areas receiving the highest votes for high policy relevance and need of further development work included non-cognitive outcomes (Modules 4 and 10), teaching and learning (Modules 2, 12 and 1), and school policies (Modules 19 and 15). These modules will be discussed in detail in the following sections. Considerable efforts have been made to include measures for those modules in the PISA 2015 field trial and main survey. Other modules will be discussed in a less detailed manner, as their content was drawn from previous cycles with little change.

Assessing non-cognitive outcomes[4]

This section summarises the conceptual foundations for high-priority Modules 10 (domain-general student behaviour and attitudes) and 4 (science-related outcomes: motivation, attitudes, beliefs) as well as those of lower-priority Modules 6 (science career) and 11 (dispositions for collaborative problem solving).

Traditionally, PISA assessed student outcomes in terms of achievement tests. Students' motivations, attitudes, beliefs and behaviours were seen as important precursors of and predictors for scholastic performance, educational attainment and labour-market success. But education policy and labour-market policy are increasingly concerned about these "non-cognitive outcomes" because they are instrumental for personal growth, individual success and the success of society as a whole.

Research has shown the predictive power of non-cognitive outcomes for success in secondary education, higher education and the workforce in general (e.g. Heckman, Stixrud and Urzua, 2006; Lindqvist and Vestman, 2011; Poropat, 2009; Richardson et al., 2012; Roberts et al., 2007). Also, professional and public debates often question the purely achievement-oriented approach of past student assessments. There is more to education than knowledge and cognitive skills; thus non-cognitive outcomes have become increasingly interesting as stand-alone outcomes in their own right. Non-cognitive dispositions are important goals, and they often function as moderators and mediators for relations of other constructs in the assessment. PISA offers a unique opportunity to investigate complex relations between non-cognitive outcomes and achievement at the individual, school and country levels.

Previous PISA cycles have focused on domain-specific student attitudes and behaviours, for instance measuring attitudes towards reading and mathematics, mathematics self-concept, or maths anxiety. Most of these scales display robust relations with student proficiency scores. This tradition is continued with Module 4 (science-related outcomes) in PISA 2015. In addition, the current framework includes a set of domain-general, non-cognitive student factors to broaden the coverage of relevant constructs, increase the policy relevance of the PISA 2015 database, and acknowledge the increased interest in non-cognitive assessments in both policy and research. Questions cover, for example, general achievement motivation. As in PISA 2012, anchoring vignettes (King and Wand, 2007) are used for better measurement to identify and correct for construct-unrelated response styles. This increases the cross-cultural comparability of the derived indices.

Figure 6.3 ■ **Measures of non-cognitive outcomes included in the PISA 2015 main survey**

Area	Science-related (Module 4)	Domain-general (Modules 6, 10, 11)
Self	**Self-efficacy**	Test anxiety Well-being in general (life satisfaction) **Well-being at school (sense of belonging)**
Interest, attitudes, and motivation	Interest in broad science topics **Enjoyment of science** **Instrumental motivation**	Achievement motivation
Beliefs and preferences	Epistemological beliefs **Environmental awareness** **Environmental optimism**	Collaboration and teamwork dispositions Career aspirations
Technology – ICT		**ICT use** Interest in ICT Perceived ICT competence Perceived Autonomy in using ICT ICT use in social interaction
Behaviour		Health: physical activities Time use: activities before/after school

Note: bold = trend measures.

Science-related outcomes (Module 4)

As science is the major domain assessed in PISA 2015, students' interest and motivation in science, technology, engineering and mathematics (STEM) subjects, plus related beliefs and behaviour are considered to be an important assessment dimension. The second column in Figure.6.3 provides an overview of the respective constructs in the broader domains of self, interest, attitudes, and motivation, beliefs and preferences.

Motivation to learn science and beliefs about oneself as a science learner are important policy and education goals in many countries. Moreover, it has been shown that they are positively associated with students' performance in science (OECD, 2007). This relationship is reciprocal: science-related beliefs and attitudes can be both a consequence and a cause of better science performance.

From a policy point of view, the shortage of skilled workers in technical and science professions – especially among women – has become a concern in recent years, and it is expected to rise in the future (European Commission, 2004, 2006; OECD, 2008). Therefore, PISA aims to identify how interested students are in science. Measures of enjoyment of science and instrumental motivation will allow for reporting trends since 2006.

In addition, environmental issues are a global concern. Threats to the environment are prominently discussed in the media, and students are challenged to understand complex environmental issues. Further, students' levels of environmental awareness and optimism affect their engagement in environmental concerns and, in turn, affect the world's global climate, the economy and society as a whole. Therefore, PISA 2015 includes two measures of environmental beliefs that were developed for PISA 2006.

The following paragraphs provide relevant research background information and the different measures included in the PISA 2015 field trial to cover these outcomes.

Self-related beliefs referring to science learning: Positive self-efficacy belief is highly related to motivation, learning behaviour, general expectations for the future and students' performance (OECD, 2007).

Motivation to learn science covers three constructs: interest in broad science topics, enjoyment of science and instrumental motivation. Motivation to learn that is based on interest and enjoyment is experienced as self-determinate and intrinsic (Krapp and Prenzel, 2011). It affects student engagement, learning activities, performance and career choices, and can be shaped by classroom instructions and parental motivation practices (Gottfried et al., 2009; Kunter, 2005; Rakoczy, Klieme

and Pauli, 2008; Ryan and Deci, 2000). In addition, instrumental motivation is an important predictor for course selection, career choices and performance (Eccles, 1994; Eccles and Wigfield, 1995; Wigfield, Eccles and Rodriguez, 1998). All three constructs have been used in previous PISA cycles, but the interest scale has been substantially revised and split into two measures.

Beliefs about science: Epistemological beliefs are closely related to students' general values of science and scientific inquiry (Fleener, 1996; Hofer and Pintrich, 2002). They include beliefs about science as an evolving and changing subject and how individuals justify knowledge (Conley et al., 2004). Epistemological beliefs have been assessed in the PISA domain of mathematics but not yet in the domain of science.

Environmental issues are a distinct area of beliefs that is covered by two constructs: environmental awareness and environmental optimism. PISA 2006 showed that students from more socio-economically advantaged backgrounds reported higher levels of awareness of environmental issues and that this construct is linked with students' science performance (OECD, 2007). In addition, students reported low levels of environmental optimism, which was negatively associated with students' performance.

Domain-general student attitudes and behaviours (Module 10)

Domain-general attitudes, beliefs and behaviours can be defined as a set of student factors or constructs that cut across curricular topics, or that are independent of curricula, but are nevertheless important for and reflective of success in education. PISA 2015 does not intend to measure global dispositional traits but behavioural tendencies and preferences that are relevant in the context of learning and that can be conceptualised as outcomes of education in general, and schooling specifically.

As shown in the right column of Figure 6.3, the content of Module 10 can be grouped into broader areas that partly parallel the science-related outcomes, widening the goal of assessment well beyond science and technology. The following paragraphs provide links to previous work at the OECD and other research, especially focusing on constructs that are new to PISA, such as well-being, health and time use.

Self-related beliefs and attitudes towards school: Generalised beliefs about one's own success or failure in academic learning have been shown to be strong predictors for further effort and success, including test scores in student assessments. PISA 2015 uses a revised (and generalised) version of a test-anxiety scale that was predictive for mathematics achievement in previous PISA cycles. In addition, a new indicator for generalised achievement motivation (e.g. "I want to be the best, whatever I do") has been introduced. A set of anchoring vignettes (King and Wand, 2007) is implemented to control for response biases and increase cross-cultural equivalence.

Subjective well-being: Subjective well-being can be defined as "good mental states, including all of the various evaluations, positive and negative, that people make of their lives and the affective reactions of people to their experiences" (OECD, 2013, p. 10). The growing recent interest in this construct among researchers and policy makers has resulted in recommendations to statistical agencies to "incorporate questions on subjective well-being in their standard surveys to capture people's life evaluations, hedonic experiences and life priorities" (Stiglitz et al., 2009: 216).

The OECD (2013) responded to this charge by providing guidelines on measuring subjective well-being. To date, 27 out of 34 OECD national statistical offices have committed to collecting at least the minimal information proposed by the OECD guidelines (the single "general life satisfaction" question), which is now included in PISA 2015. The guidelines suggest that it is appropriate to collect such information from 15-year-olds, and even younger students, because the evidence suggests that they are "capable of responding effectively to subjective well-being questions from as young as age 11 with respect to measures of life evaluation and affective state" (OECD, 2013: 152). To assess well-being specifically within the context of school, PISA 2015 uses a question that has been labelled "sense of belonging" in previous cycles.

ICT: ICT-related behavioural characteristics and motivational attributes can be regarded as domain-general student outcomes. Since ICT subsumes a broad range of devices, it may play a role across all education domains. Following the OECD's DeSeCo project and the 21st-Century Skills Initiative, students should exhibit general skills related to information, media and technology above and beyond the traditional core subjects (OECD, 2005; Partnership for 21st Century Skills, 2008). PISA 2015 assesses students' interest in ICT, (self-determined) practices of ICT use, their perceived competence and autonomy in using ICT, and a specific question on use of social media within the ICT Familiarity Questionnaire.

Health: This area addresses the practice of healthy behaviours, particularly behaviour pertaining to regular exercise, and health-related lifestyle choices. Research has shown that poverty and low socio-economic status is associated with poor health outcomes (Spurrier et al., 2003). Research has also shown that physical activity can improve academic performance due to the activity itself and to the consequent reduction of more passive activities (Salmon et al., 2005).

Physical exercise can be (positively as well as negatively) influenced by teacher behaviours and school practices. The approach to measuring physical activity in PISA 2015 is adapted from and closely aligned with the Global School-based Student Health Survey of the World Health Organization. That survey is conducted among adolescents of roughly the same age as the PISA population. Relating these data to PISA measures of student and school background will help to develop an understanding of the issues of equity and health in education.

Career aspirations (Module 6)

The PISA 2015 Student Questionnaire includes two questions about career aspirations that were used in 2006. They ask about the highest education level the student expects to complete and the job he or she expects to have at the age of 30.

Non-cognitive outcomes related to collaborative problem solving (Module 11)

In order to cover dispositions related to the new domain of assessment introduced in PISA 2015, namely Collaborative Problem Solving, a set of items on valuing team work, co-operating, guiding others and negotiating was developed based on research by Wang and colleagues (2009). The Teacher Questionnaire covers types of activities and grouping, and rewards for team work from another perspective.

Assessing teaching and learning processes[5]

This section summarises the conceptual foundations for high-priority Modules 2 (science teaching practices), 12 (learning time and curriculum) and 1 (teacher qualifications and professional knowledge) as well as those of lower-priority Module 5 (out-of-school science experience).

Teaching and learning are at the heart of schooling. Most cognitive and non-cognitive, curricular and cross-curricular goals of school education are achieved – or not – by the way students and teachers interact in classrooms. While teaching is the core process in schools, the curriculum determines its content, and professional teachers implement the curriculum, orchestrate learning activities, and thus arrange for quality learning time.

PISA has been designed as a yield study, assessing life-skills and broad areas of literacy rather than curricular domains, sampling a birth cohort rather than a grade level or intact classrooms. Thus, it might be questioned why this programme should address teaching and learning processes at all. However, there is ample evidence that teaching and learning activities are the best predictors of student competencies, whatever their character might be. So, if PISA is to inform education policy making at the system and the school levels, it must cover this important area.

Clearly, the PISA study should focus on more general and internationally comparable constructs rather than fine-grained content matter. Therefore, Module 2 describes science education by broad lists of teaching and learning activities, covering both inquiry-based teaching (which was assessed in PISA 2006) and teacher-directed practices. In addition, general dimensions of teaching quality, such as classroom disciplinary climate, teacher support, feedback and adaptability, are applied to science education. Module 12 covers learning time – including non-mandatory, additional instruction within and out of school – as well as the science curriculum. In addition, the teaching force is described in terms of initial education, beliefs and professional development (Module 1).

Science teaching practices (Module 2)

According to the PISA approach to scientific literacy, the main task of science teaching is to foster students' capacity to explain phenomena scientifically, understand scientific enquiry and interpret scientific evidence. The key topic of the framework outlined below is the extent to which schools are mastering this task.

A number of processes at the classroom level have been found to be relevant for effectiveness in science education. In this framework, both domain-specific instructional approaches and activities, and more general dimensions of instructional quality are combined, as they are equally suited to support learning activities and describe processes on the classroom level. However, in PISA 2015, all questions about teaching and learning activities are framed within the context of school science, sometimes even referring to one specific course. The aim is to describe science teaching in the classroom by country-specific profiles of teaching practices, and to investigate their relation to students' outcomes.

Analyses based on PISA 2006 show that a student's outcome can be predicted by different profiles of practices in teaching (Kobarg et al., 2011). While some teaching patterns are related to high performance, others are related to high student interest and motivation. The results indicate that the items and scales for science teaching practices are applicable to in-depth descriptions of science teaching in the classroom. Moreover, a comparison of the patterns allows for detailed analyses of both students' science performance and students' interest in science topics across countries (Kobarg et al., 2011; Prenzel, Seidel and Kobarg, 2012). The teaching-practices items are developed and chosen in order to discriminate between different patterns of teaching.

Teaching and learning activities: Research has shown that inquiry-based teaching practices, which play a significant role in science education, have a positive effect on student learning, particularly students' engagement in the cognitive dimensions of inquiry and teacher-led inquiry activities (Furtak et al., 2012). Inquiry-based instruction seems not only to improve achievement (Blanchard et al., 2010), but also attitudes towards the subject and transferable critical thinking skills (Hattie, 2009).

There is a renewed interest in embedding science teaching and learning in contexts that are real and meaningful for learners (Fensham, 2009; King and Stephen, 2012). Scientific argumentation as a central goal of science education (Osborne, 2012) needs classroom situations with sufficient opportunities for social interaction. Instruction that emphasises students' active thinking and drawing conclusions from data seems to be particularly beneficial for students' development (Minner, Levy and Century, 2010). According to these findings and the analysis of the PISA 2006 items (Kobarg et al., 2011, Taylor, Stuhlsatz and Bybee, 2009), a subset of nine items from PISA 2006 is used for this scale: six of the items are unchanged from 2006; three were slightly revised.

In addition to the inquiry-based teaching practices, teacher-directed teaching and learning activities in science focus on classroom management activities and teaching methods, and broaden the perspective of domain-specific practices. The purpose is to obtain student-reported information about their actions in school science lessons and to get a realistic picture of what is going on in science classes – including classes with little inquiry-based learning.

The student perspective on science teaching is complemented by the Teacher Questionnaire, for those countries that participate in this option. Science teachers are asked to describe their teaching practices through a parallel questionnaire that also focuses on teacher-directed teaching and learning activities in science lessons, and a selected set of inquiry-based activities. Both perspectives may be combined and compared on the school level.

Dimensions of teaching quality: Several classroom studies confirm the impact of three basic dimensions of instructional quality on students' cognitive and motivational development: clear, well-structured classroom management; supportive, student-oriented classroom climate; and cognitive activation with challenging content (Klieme, Pauli and Reusser, 2009). In previous PISA cycles, the first two dimensions were covered by questions about classroom disciplinary climate and teacher support, respectively. As a school-climate variable, the purpose of the disciplinary climate question is to gather information on the structure and efficiency of classroom management, which can be seen as a prerequisite for student learning. The teacher-support question measures how often the teacher helps students with their learning (OECD, 2004).

Research has shown that the scale is positively related to students' interest (Vieluf, Lee and Kyllonen, 2009). Concerning a measure of cognitive activation, it is assumed that the level of cognitive challenge is determined by the type of problem and the way it is presented in the lesson. Therefore, "inquiry-based science education" serves as an indicator of cognitive activation in PISA 2015. In addition to these three dimensions, PISA 2015 includes a measure of adaptability in teaching, as perceived by students.

Learning time and curriculum (Module 12), including out-of-school science experience (Module 5)

The learning time and curriculum to which students are exposed in the course of their education are closely related to student outcomes (e.g. Schmidt and Maier, 2009; Abedi et al., 2006; Scherff and Piazza, 2008).

Learning time has proven to be a central factor in student learning and achievement (Gándara et al., 2003; Patall, Cooper and Allen, 2010; Scheerens and Bosker, 1997; Seidel and Shavelson, 2007). Such positive relationships were replicated in international comparative research, pointing to the cross-cultural comparability of the construct and its effects (e.g. OECD, 2011; Martin et al., 2008; Schmidt et al., 2001). Yet although there is an overall positive relation between learning time and achievement, there are large differences within and across countries and among different groups of students and schools (Ghuman and Lloyd, 2010; OECD, 2011).

Overall it is important to distinguish among learning time that is provided by the school system, realised or implemented by the school and the teacher in the classroom, and used by the students. On this path from "gross" learning time as allocated in system-level policies to student "time-on-(the right) task", many factors, at different levels (school, classroom and student) reduce available learning time to varying degrees across countries (Gillies and Quijada, 2008; Benavot, 2004). Differences in that reduction among various subgroups of students can indicate equity – or lack thereof – in education opportunities, because research shows that relations with outcomes are stronger when learning time is more narrowly defined (e.g. time-on-task instead of allocated learning time). Therefore, PISA 2015 intends to apply a broader view of learning time (Abadzi, 2009; Berliner, 1990; Millot and Lane, 2002).

At the school level, PISA 2015 assesses provided learning time (PT), but there is loss of time due to such factors as local festivities, teachers' strikes, illness or other teacher absenteeism (see Ghuman and Lloyd, 2010; Chaudhury et al., 2006). Another proportion of time is then used in the classroom, resulting in realised learning time (rt). Time loss at the classroom level is most commonly due to non-teaching activities like classroom management, collecting homework or waiting time (e.g. MacKay, 2009; The PROBE Team, 1999), which leaves realised learning time as the fraction of time during which a class is taught.[6] The proportion of realised learning time during which a student learns course content is engaged learning time (ET). This excludes periods in which a student does not attend class due to illness, truancy, being late or being physically present but mentally disengaged. Engaged learning time is the only time during which students actually learn. Figure 6.4 provides a summary of these time-related constructs, how they are defined and how they may be estimated.

Moreover, it has been shown that next to the absolute amount of time available for learning, students' time-use patterns relate to success variables and can help explain associations between student background variables (such as socio-economic status) and performance variables (such as mediator variables, see Porterfield and Winkler, 2007). In the PISA 2015 main survey, students' time use before and after school is assessed with a set of newly designed questions that was developed in parallel to Kahneman's et al. (2004) "Day-reconstruction method".

Figure 6.4 ▪ **Assessment of learning time and loss of learning time in PISA 2015**

		Student Questionnaire	School Questionnaire	
Use	**Student**	+ Additional instruction and study (time use) - Truancy		Engaged time (ET) = RT – student absenteeism, truancy, mentally disengaged time
Provision	**Classroom**	- Disciplinary climate and loss in science classes		Realised learning time (RT) = PT – loss due to classroom management, assessment time, waiting time, etc.
Provision	**School**	+ Amount of school learning time + Number and type of science classes	- Loss on school level	Provided learning time (PT) = AT – loss due to weather, holidays, teacher absenteeism, etc.

In addition to learning times for mandatory schooling, other in-school and out-of-school learning activities are also taken into account. PISA 2015 attempts to identify additional learning time in a cross-culturally valid way, incorporating, for example, different formats, location, content and purposes. Information from the School and Student Questionnaires, and from the optional Educational Career Questionnaire can be combined to get the full picture. Similarly, information on extracurricular learning activities, time use before and after school every day, and science-related experiences is gathered from students, from parents in the optional Parent Questionnaire, and from school leaders in the School Questionnaire.

Curriculum: There may be great differences between the curriculum designed at the system level, the curriculum delivered by the teacher or in the textbook, and the curriculum as understood by students. For the major domain of PISA 2015, "science", differences in the curriculum are particularly large across tracks, grades, schools and countries (Schmidt et al., 2001; Martin et al., 2008). In order to examine some of this variety, the optional PISA 2015 Science Teacher Questionnaire asks about the content of a school's science curriculum and how it is communicated to the parents.

Teacher qualification and knowledge/beliefs (Module 1)

Many studies show a clear link between teacher-related factors and student learning. In addition to teachers' professional behaviour within the classroom (see above), the age and education level of the teaching force, teachers' initial education and qualifications, their individual beliefs and competencies, and their professional practices on the school level, such as collaboration and professional development, have been core topics in education policy.

Some basic information on these topics will be available from the PISA 2015 School Questionnaire, while the optional Teacher Questionnaire, which is partly based on instruments previously established in the OECD Teaching and Learning International Survey (TALIS), features additional constructs, both science-specific and domain-general (Figure 6.5). This instrument is new to PISA, although national instruments have been added to the PISA design in Germany and Ireland in previous cycles, with broad support from the teaching force. Other large-scale studies, such as Trends in International Mathematics and Science Study (TIMSS), have implemented teacher questionnaires without any loss in participation. Thus, the new optional instrument will finally give teachers a voice in PISA.

Figure 6.5 ▪ **Teacher-related measures in the PISA 2015 field trial**

	Science-related	General
Background	Gender, age, employment status, job experience, subjects studied	
Initial education	Goal of first qualification, type of teacher education and training programme (if attended), mode of qualification Number of teachers by education level (ScQ)	
	Science-related content Number of science teachers by level of qualification (ScQ)	
Professional development	Participation in different type of activities Obligation amount of participation, school policies (ScQ)	
	Collaboration Science-related content	Co-operation General content
Beliefs	Self-efficacy (related to science content and teaching science)	Job satisfaction

Note: If not indicated otherwise, constructs are included in the optional PISA 2015 Teacher Questionnaire.

Across these topics, a distinction is adapted that Shulman (1985) suggested for research on teachers: teachers' beliefs and activities can be related either to the subject matter taught, its conceptual foundations, basic ideas, etc. (content), to teaching and learning the subject matter, including issues of student understanding, teaching practices, assessment procedures, etc. (pedagogical content), or to general concepts, such as classroom management (pedagogy).

Shulman's model has been most influential in research on teachers (e.g. Hill, Rowan and Ball, 2005; Baumert et al., 2010; Bloemeke et al., 2012). In line with this research, PISA 2015 identifies content, pedagogical content, and/or pedagogy as the foci of teacher-related constructs, including initial education and professional development. There is no attempt to measure teachers' knowledge.

Teachers' background and initial education: Understanding the multiple pathways leading to the teaching profession, including mid-life career changes, is important for education policy because there is a growing need to recruit teachers from non-traditional backgrounds. For these people, but also for novice teachers with traditional training, the induction stage is important (Portner, 2005). Teacher retention is another concern in many countries (Ingersoll and Perda, 2010). In addition to formal qualifications (tertiary/secondary education certificates and academic degrees), a major in the subject being taught, the type of teacher education and training programmes attended, and professional experience (i.e. years having taught science at school), PISA 2015 asks teachers about the representation of the three foci in their initial education.

Professional development and collaboration: Professional development refers to any activity that equips teachers with the tools and resources necessary to provide quality instruction. It includes school-based programmes as well as networking, coaching, seminars or other types of training activities that foster in-service learning and thus promote the professionalisation of teaching. Even though professional development is generally regarded as crucial for improving

teaching and student achievement, Sykes referred to the ineffectiveness of common trainings as "the most serious unsolved problem for policy and practice" (Sykes, 1996: 465). However, more recent studies report positive effects on teaching practices and classroom climate (Cuevas et al., 2005; Desimone et al., 2002; Jeanpierre, Oberhause and Freeman, 2005; Supovitz and Turner, 2000; Timperley et al., 2007), as well as on student achievement (e.g. McDowall et al., 2007; Shayer and Adhami, 2007). This apparent inconsistency may be partly resolved by accounting for different features of the programmes examined. Summarising previous studies, Buczynski and Hansen describe ineffective programmes as being "too conventionally taught, too top-down, and too isolated from school and classroom realities to have much impact on practice" (Buczynski and Hansen, 2010: 600).

As early as in the 1980s, scholars indicated the benefits of supportive networks for teachers (e.g. Darling-Hammond, 1984; Rosenholtz, 1989; Bryk and Driscoll, 1988). In the 1990s the idea of "professional learning communities" emerged. This notion refers to groups of teachers who co-operatively reflect and improve their professional practices (Hord, 1997). Research on professional learning communities is still limited, but there is some indication of positive effects on education processes and outcomes (e.g. Lomos, Hofman and Bosker, 2011). In China, for example, teachers are often organised in groups that work together studying national guidelines and defining teaching goals, that co-operate for preparing and improving teaching, and that organise observation visits to provide colleagues with feedback and involve teachers in out-of-school activities (Paine and Ma, 1993). Similarly, in Japan "lesson studies" are common practice among teachers (Stigler and Hiebert, 1999). TALIS further suggests that the pattern of activities also varies between countries (Vieluf et al., 2011).

The PISA School Questionnaire in 2000 and 2012 included a question about the proportion of teachers who had recently (within the previous three months) participated in some kind of professional development activity. In 2012, the same question was asked with a focus on mathematics teachers. However, this information did not show any substantial relation to student outcomes. Therefore, PISA 2015 intended to enhance measurement of professional development by adapting questions from TALIS and other sources (e.g. Steinert et al., 2006).

Professional beliefs: PISA uses a measure of job satisfaction also used in TALIS. Science teachers are asked to report self-efficacy beliefs regarding science content and how it is taught.

Teachers' morale and commitment was assessed in PISA cycles 2000, 2003 and 2012 in the School Questionnaire that was completed by the principal (or some other member of the school management team), aiming to assess attitudes among teaching staff. These measures are included in PISA 2015 in Module 13, "school climate".

The main level of analysis for data gathered in the optional Teacher Questionnaire is the school level. No weighting for individual teacher responses are available. All data from the Teacher Questionnaire are thus treated as school variables.

Assessing school policies and governance[7]

This section summarises the conceptual foundations of high-priority Module 19 (assessment, evaluation and accountability) as well as those of lower-priority Modules 3 (school-level learning environment for science) and 13-18.

Assessment, evaluation, and accountability (Module 19)

Assessing students and evaluating schools[8] is a common practice in most countries. Since the 1980s, policy instruments, such as performance standards, standard-based assessment, annual reports on student progress and school inspectorates, have been promoted and implemented across continents. Reporting and sharing data from assessments and evaluations with different stakeholders provides multiple opportunities for monitoring, feedback and improvement.

In recent years, there has been a growing interest in the use of assessment and evaluation results through feedback to students, parents, teachers and schools as one of the most powerful tools for quality management and improvement (OECD 2010d: 76). Accountability systems based on these instruments are increasingly common in OECD countries (Scheerens, 2002: 36). Accountability is often linked to market-oriented reforms. Rewards/penalties for good/poor assessment and evaluation results are said to change behaviours in ways that improve student achievement (Wößmann et al., 2009). However, there are huge differences in assessment and evaluation practices and purposes.[9]

Previous PISA cycles have covered aspects of assessment, evaluation and accountability in the School Questionnaire, with a strong focus on the use of standardised tests. In PISA 2015, this module asks both about standardised and less standardised practices. Internal and external evaluations address different purposes and consequences and will be dealt with separately. Teacher evaluation is also addressed as a means of quality management.

Formative assessment and feedback are increasingly popular in research and teaching practice. These types of assessment and evaluation differ with respect to their respective purposes and criteria, practices, use and consequences (Pellegrino, Chudowsky and Glaser, 2001; Scriven, 2003; Wilson, 2004) (see Figure 6.6). These aspects are covered in the PISA 2015 questionnaires as much as possible.

In the following section, relevant research on school evaluation and student assessment is summarised to provide the rationale for questionnaire development in PISA 2015.

Evaluation: The evaluation of schools is used as a means of assuring transparency, judging systems, programmes, educational resources and processes, and also guiding school development (Faubert, 2009). Evaluation criteria need to be defined and applied from the viewpoints of different stakeholders (Sanders and Davidson, 2003).

Evaluation can be either external or internal (Berkemeyer and Müller, 2010). It is called external evaluation if the process is controlled and headed by an external body and the school does not define the areas that are judged. An evaluation is called internal if it is part of a process controlled by the school and in which the school defines which areas are judged. The evaluation may be conducted by members of the school (self-evaluation) or by persons/institutions commissioned by the school. Different evaluation practices generally coexist and benefit from each other (Ryan, Chandler and Samuels, 2007).

External evaluation can expand the scope of internal evaluation, validate results and implement standards or goals; internal evaluation can improve the interpretation and increase the use of external evaluation results (Nevo, 2002). However, improvement of schools seems to be more likely when an internal evaluation is applied, compared to external evaluation. Therefore, processes and outcomes of evaluation might differ, depending on whether the evaluation is internal or external. Moreover, country- and school-specific factors may influence the implementation of evaluations as well as the conclusions and effects for schools. In many countries, individual evaluations of teachers and principals, separate from school-wide evaluation, are also common (Faubert, 2009; Santiago and Benavides, 2009); they are treated here as a separate type of evaluation.

Figure 6.6 ▪ **Measures in PISA 2015 related to assessment, evaluation and accountability**

<table>
<tr><th></th><th>External evaluation</th><th>Teacher evaluation</th><th>Internal evaluation</th><th>Formative assessment</th></tr>
<tr><td rowspan="2">Purpose and criteria</td><td colspan="4">General assessment practice (ScQ)
Purpose of assessment results (ScQ)</td></tr>
<tr><td colspan="2">Evaluation policies (ScQ)</td><td></td><td>Teacher's grading (TQG)</td></tr>
<tr><td>Practices</td><td></td><td>Teacher-evaluation methods (ScQ)</td><td></td><td>Classroom-assessment instruments (TQG/TALIS)</td></tr>
<tr><td>Use and consequences</td><td>Processes of external evaluation (ScQ)
Use of achievement data for accountability (ScQ)</td><td></td><td>Consequences of internal evaluation (ScQ)</td><td>Feedback: student perception (StQ)
Adaptation of instruction (StQ)</td></tr>
</table>

Results of evaluations may be used in a formative way (e.g. to guide the analysis and improvement of processes) or in a more summative way (e.g. for accountability). Formative evaluation aims at closing the gap between the as-is state and the target state. Here, teaching and school-based processes are to be guided to a predetermined goal. Summative evaluation focuses on student outcomes and encourages schools to meet specific standards. Formative evaluation has turned out to be more effective in school improvement than summative evaluation (Creemers and Kyriakides, 2008). Effects or consequences of evaluation may differ, depending on the focus of evaluation, the procedure chosen for the evaluation, or a school's goals and priorities.

Assessment: Communication and clarification of achievement goals within schools is essential in students' learning (Brookhart, 2007; Stiggins, 2007). National standards that have emerged in recent years define what students should know (Koeppen et al., 2008; Shepard, 2006). These education standards directly shape school policies and classroom instruction by urging schools and teachers to communicate specific aims leading to a shared understanding. To check whether these goals are met, schools follow a given assessment practice or define their own. This can be implemented in the classroom learning process by more or less standardised tests and oral examinations developed by the teacher.

In addition, mandatory and non-mandatory standardised and externally developed tests verify and compare student outcomes at the classroom, school, district, state or international level (Shepard, 2006). Irrespective of the purpose and the stakeholder administering the assessment, a test must fulfil a number of quality criteria (Scheerens, Glas and Thomas, 2003). In general, standardised tests are more reliable measures, but may be less aligned with the school curriculum, and vice versa for teacher-made assessments.

The distinction between formative and summative, internal and external approaches also holds for student assessment. In its summarising function, assessment takes place in order to grade, certify or record progress. Summative assessment, whether external or internal, thus indicates and monitors standards, but may also raise standards by encouraging students, as well as teachers and schools, to put more effort into their work (Harlen and Deakin Crick, 2002). On the other hand, summative assessment might lead to lower self-esteem and diminished effort of students at risk, and therefore can increase the gap between lower- and higher-achieving students (Black and Wiliam, 2004).

Another negative aspect of assessment may arise when teaching solely focuses on answering questions, rather than on developing skills and knowledge (Harlen and Deakin Crick, 2002). Grading is arguably the most prevalent type of classroom assessment, and an essential aspect of effective teaching (McMillan, 2001; Guskey, 2007). Grades have been shown to be unreliable and of limited validity, but there is very little comparative research on grading practices in different countries.

Formative assessment can be a significant source of improvement in student learning processes (e.g. Shepard, 2006; Black and Wiliam, 2004; McMillan, 2007; OECD, 2006b). Especially for low achievers, formative assessment can lead to sizable gains in student achievement (Abrams, 2007). Formative assessment and reciprocal feedback might not just be useful for students, but also for teachers, helping them to adapt their instruction to their students' needs. However, there is large variation in the implementation and impact of formative assessment practices (e.g. Kingston and Nash, 2011; Shute, 2008; Hattie and Timperley, 2007; Black and Wiliam, 1998). Therefore, it is worthwhile to study cross-national variation in formative assessment practices through PISA 2015.

The School Questionnaire for PISA 2015 includes several questions on general assessment practices and results, external evaluation and teacher evaluation that have been used in previous cycles, to report trends. However, in line with the research cited above, internal school evaluation and formative, classroom-based assessments are given more weight in PISA 2015 than in previous cycles.

Other school policies and approaches to educational governance

During the past two decades, research on educational effectiveness has largely been concerned with the impact of school-level factors on students' learning. Studies show that school qualities have effects on student progress, with variation in schools appearing to affect students' behaviour. It has been asserted that the environment at the school level can influence the behaviour of teachers and students and thus – mostly indirectly – their consequent success in teaching and in learning. Both "soft" factors, such as school climate and parental involvement, and "hard" factors, such as school management activities and allocation policies, vary and are related to student outcomes within and across countries.

School climate (Module 13): School climate encompasses shared norms and values, the quality of relationships and the general atmosphere of a school. An academic focus – a general consensus about the mission of the school and the value of education, shared by school leaders, staff, and parents – influences the norms in student peer groups and facilitates learning. In addition, an orderly learning atmosphere maximises the use of learning time. By contrast, disrespectfulness and an unruly environment are counterproductive for teachers and students alike and distract from the school's actual mission. As in previous PISA assessments, indicators for school climate is gathered through the School Questionnaire ("behaviour affecting school climate").

PISA 2015 adds two new measures covering aspects of school climate that are often hidden, but nevertheless highly important from both a pedagogical and a policy point of view: bullying by peers and unfair treatment by teachers, as perceived by the students. Bullying has been identified as an important factor of school culture (Ertesvag and Roland, 2015) and school climate (Wang, Berry and Swearer, 2013) that is relevant across cultures (Smith et al., 2002).

Parental involvement (Module 14): Over the past years, the involvement of parents in education has gained importance in the education debate, and to some extent it has also become relevant for education policy. Parents are not only an important audience, but powerful stakeholders in education. Thus information on parents' opinions and engagement is highly valuable for large-scale assessments like PISA. Parental involvement in education has been assessed in PISA

since 2006 when the Parent Questionnaire, which directly addresses the parents of PISA students, was administered for the first time. For PISA 2015, specific aspects of parental involvement are added to the School Questionnaire (for parent-school communication and collaboration), and to the Student Questionnaire (for parental support in learning). In particular, four items on parental support are asked in parallel of students and of their parents, so that perceptions can be compared on an individual level.

Leadership and school management (Module 15): School principals play a key role in school management. They can shape teachers' professional development, define the school's education goals, ensure that instructional practice is directed towards achieving these goals, suggest modifications to improve teaching practices, and help solve problems that may arise within the classroom or among teachers. The PISA measure of educational leadership saw a major revision in the 2012 study. This work has been reviewed, and the leadership scale could be considerably shortened. In addition, the new optional Teacher Questionnaire gathers information on transformational leadership from teachers as well, because research has shown that the teachers' perspective on leadership can differ from the positions held by school administrators.

Resources (Module 16): Information on school type (public vs. private) and class size has always been included in the School Questionnaire. In addition to these trend questions, PISA 2015 allows for distinguishing between types of private schools (religious/denominational, not-for-profit, for-profit). All PISA cycles so far have included a question on the degree of problems a school experiences due to missing resources. The different approaches over time are systematised and implemented in one coherent question in the School Questionnaire.

Locus of decision making (Module 17): Education systems have been classified by the amount of control that is left to the school (i.e. school board, staff and school leaders) when decisions on admission, curriculum, allocation of resources and personnel have to be made. These indicators are based on questions in the School Questionnaire that are left unchanged to allow for trend reporting.

Allocation, selection, choice and grade repetition (Module 18): The way students are channelled into education pathways, schools, tracks or courses is a core issue of educational governance ("stratification"). On the school level, selection and allocation procedures are important aspects of school organisation. Highly selective schools provide a learning environment that may differ from the environment offered by more comprehensive schools. For all those reasons, appropriate trend questions answered by school administrators and parents have been retained.

School learning environment for science (Module 3): Conceptually, this module overlaps to a considerable degree with other modules dealing with school-level factors, such as Module 12 (Learning time and curriculum), Module 15 (School leadership and management), and Module 19 (Assessment, evaluation and accountability). In addition to those, there are questions addressing the size and qualification level of the science teaching staff, and concerning available resources, such as laboratories and equipment for hands-on student activities.

Assessing student background (Modules 7–9)[10]

This section covers three modules that were given lower priority by the PISA Governing Board: Module 7 (student socio-economic status, family and home background), Module 8 (ethnicity and migration) and Module 9 (educational pathways in early childhood). Nevertheless, these topics, and Module 7 in particular, are important, because they contain the basic information needed for calculation of the PISA index of economic, social and cultural status (ESCS).

Student socio-economic status, family and home background (Module 7): In order to compare equity related to social and ethnic factors across PISA cycles, PISA 2015 keeps measures of socio-economic status and other background variables basically unchanged. However, some minor changes have become necessary. Due to extensive development in the ICT sector, for example, questions on technology equipment in the student's home are slightly outdated. Thus, the measures of home possessions have been updated to ensure better coverage of within and across country variation of home possessions. These changes are expected not to have any effect on the important trend measures in this module.

Ethnicity and migration (Module 8): Linguistic and cultural diversity are basic facts of life in most regions of the world. Many nations are home to several subpopulations with different languages and cultures. International migration perpetuates this diversity. In OECD countries, first- and second-generation immigrant students currently comprise 10% to 20% of the student population (OECD, 2010). At the same time, students from ethnic minority groups and immigrant students often face particular challenges. In a number of education systems, immigrant students perform at significantly lower levels than their non-immigrant peers in key school subjects (Stanat and Christensen, 2006); both groups are often faced with overt or covert discrimination with potentially detrimental consequences for their psychological development and well-being. Thus, providing students from different linguistic and cultural backgrounds with equal opportunities is often considered

one of the central challenges for education systems in the 21st century. Due to cultural concerns in some countries and time restrictions in the Student Questionnaires, however, PISA 2015 retains the questions on immigration and language background that have been used in previous cycles. This information is slightly enhanced by asking principals to estimate the percentage of minority students (language minorities, socio-economically disadvantaged and special needs) among 15-year-old students in their school.

Education pathways in early childhood (Module 9): When children enter primary school, they already differ in their language, pre-reading and early numeracy skills, and these differences are often maintained later in life. Promoting school readiness and better adjustment to school is hypothesised to be an efficient means of raising the achievement levels of all children, but especially of those children who lack parental support or who grow up in disadvantaged circumstances. It has been argued that investing in early education programmes will have large, long-term monetary and non-monetary benefits (Heckman, 2006). The importance of pre-school quality has been acknowledged and analysed by OECD reporting as well.

According to UNESCO (2006), early childhood care and education (ECCE) is defined as "programmes that, in addition to providing children with care, offer a structured and purposeful set of learning activities either in formal institutions (pre-primary) or as part of a non-formal child development programme". The focus of the internationally comparable statistics, International Standard Classification of Education Level 0 (ISCED 0), is much narrower. Currently, at least four strands of research support the relevance of applying a broader definition of ECCE than focusing on ISCED 0 alone: brain research, studies on domain-specific development and support, evaluation studies of model programmes, and longitudinal large-scale studies all rely on the broader definition of ECCE. Thus, conclusions about the importance of early child care should be drawn with ECCE and not with ISCED 0 in mind.

However, when evaluating the body of research it becomes obvious that, in fact, a number of characteristics of the kind of ECCE provided seem to determine whether benefits can be observed or not, and whether these benefits disappear or persist. Students' opportunities to learn in early childhood are best assessed in terms of curriculum, quantity and quality of early childhood learning experiences. For example, one of the best sources available, the British Effective Provision of Pre-School Education (EPPE) Study, did find short-term effects showing that pre-school attendance was beneficial for cognitive and socio-emotional development, particularly for children from disadvantaged backgrounds. However, in the long term only those children who attended a high-quality pre-school centre showed persistent beneficial pre-school effects (e.g. Sammons et al., 2008; Sylva et al., 2011a; see also Valenti and Tracey, 2009). A certain degree of intensity in terms of hours per week/months seems to be a precondition for beneficial effects of attendance at pre-school programmes (Logan et al., 2011; Sylva et al., 2011b).

Thus, asking about early education experience in PISA only makes sense if specific aspects of programme duration, quality and curriculum can be retrieved retrospectively, which is more than unlikely (Fivush and Hamond, 1990; Markowitsch and Welzer, 2009). As a consequence, PISA 2015, while keeping a short question on ISCED 0 attendance in the Student Questionnaire, includes a series of questions in the Parent Questionnaire, expecting parents to be a more reliable source of information. Those countries that distribute the optional Parent Questionnaire will acquire information on basic characteristics of the early childhood education and care arrangements of PISA participants, and reasons for attending or not attending early childhood education and care.

Notes

1. In the past, other, more technical notions of "core" have been used in the PISA Questionnaire Design. One approach used "core" to denote a set of variables in the Student Questionnaires that are measured for all students in a given PISA cycle – even in cases where different booklets are assigned to them. Another approach defined "core" as the set of variables used for imputing plausible values of test scores. A third approach referred to "core" as the set of domain-general variables, i.e. those not related to the major domain of assessment. Please note that in contrast to those definitions, this framework identifies "core content" as a set of conceptual constructs that defines the most basic context assessment necessary in PISA. This set of constructs (has been and) should therefore be included in all PISA cycles, albeit in some cases adapted to the major domain.

2. From a technical point of view, it is also important to note that this stable set of background variables guarantees a strong set of conditioning variables used to impute measures of student proficiencies, as explained in the PISA Technical Reports (e.g. OECD 2014: 146).

3. With the exception of optional material, such as the Parent Questionnaire, the ICT Familiarity Questionnaire, the Educational Career Questionnaire and the PISA 2000 Questionnaire on cross-curricular competencies.

4. This section is based on working papers submitted by Anja Schiepe-Tiska, Christine Sälzer and Manfred Prenzel for Module 4, Jonas Bertling and Patrick Kyllonen for Module 10. Module 11 was developed in co-operation with Core 1 and the Collaborative Problem Solving Expert Group chaired by Art Graesser.

5. This section is based on working papers submitted by Katharina Müller, Manfred Prenzel and Tina Seidel for Module 2, Susanne Kuger for Module 12, Eckhard Klieme, Franz Klingebiel and Svenja Vieluf for Module 1.

6. At least at this level, "time loss" refers to diminished learning time that focuses on curricular content and therefore on domain-specific cognitive outcomes. More overarching goals of education, such as self-regulation, interest or social competencies, might very well be stimulated during "lost" time periods.

7. This section is based on working papers submitted by Sonja Bayer, Eckhard Klieme and Nina Jude for Module 19, Leonidas Kyriakides for Module 3, Silke Hertel, Nadine Zeidler and Nina Jude for Module 14 (Parental Involvement) and Bieke de Fraine for Module 15 (School Management).

8. The terms "evaluation" and "assessment" are defined quite differently in the literature. Sometimes they are even treated as synonyms. In this section the definition used conforms to that used in current OECD literature (see e.g. Rosenkvist, 2010). The term evaluation or school evaluation is used for processes on the school and system levels. Evaluators collect evidence to judge systems, education programmes, policies and practices. This may include an evaluation of individual performance among professionals, such as teacher evaluation. Assessment or student assessment, on the other hand, directly refers to student performance or student learning processes (see also Harlen, 2007). Notably, there is a strong link between assessment and evaluation. For instance, results from student assessments may be used for school evaluation purposes.

9. See OECD's country reviews of assessment and evaluation www.oecd.org/education/preschoolandschool/oecdreviewonevaluationandassessmentframeworksforimprovingschooloutcomescountryreviews.htm.

10. This section is based on working papers submitted by Wolfram Schulz for Module 7, Svenja Vieluf for Module 8, Susanne Kuger and Hans-Günter Roßbach for Module 9.

References

Abadzi, H. (2009), "Instructional time loss in developing countries: concepts, measurement, and implications", *The World Bank Research Observer*, Vol. 24/2, pp. 267-290.

Abedi, J. et al. (2006), *English Language Learners and Math Achievement: A Study of Opportunity to Learn and Language Accommodation*, CSE Tech. Rep. No. 702, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles, CA.

Abrams, L.M. (2007), "Implications of high-stakes testing for the use of formative classroom assessment", in J.H. McMillan (ed.), *Formative Classroom Assessment: Theory into Practice*, Teachers College, Columbia University, New York, pp. 79-98.

Ainley, M. and **J. Ainley** (2011a), "Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science", *Contemporary Educational Psychology*, Vol. 36/1, pp. 4-12.

Ainley, M. and **J. Ainley** (2011b), "A cultural perspective on the structure of student interest in science", *International Journal of Science Education*, Vol. 33/1, pp. 51-71.

Almlund, M. et al. (2011), "Personality psychology and economics", *IZA Discussion Papers 5500,* Institute for the Study of Labour (IZA), Bonn.

Basl, J. (2011), "Effect of school on interest in natural sciences: A comparison of the Czech Republic, Germany, Finland, and Norway based on PISA 2006", *International Journal of Science Education*, Vol. 33/1, pp. 145-157.

Baumert, J. et al. (2010), "Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress", *American Educational Research Journal,* Vol. 47/1, pp. 133-180.

Benavot, A. (2004), "A global study of intended instructional time and official school curricula, 1980-2000", background paper commissioned by the International Bureau of Education for the *UNESCO-EFA Global Monitoring Report 2005: The Quality Imperative*, IBE, Geneva.

Berkemeyer, N. and **S. Müller** (2010), "Schulinterne Evaluation – nur ein Instrument zur Selbststeuerung von Schulen?" in H. Altrichter and K. Maag Merki (eds.), *Handbuch neue Steuerung im Schulsystem*, VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 195-218.

Berliner, D.C. (1990), "What's all the fuss about instructional time?" in M. BenPeretz and R. Bromme (eds.), *Life in Classrooms,* reissued with a new introduction, Teachers College Press, New York, pp. 3-35.

Black, P. and **D. Wiliam** (2004), "Classroom assessment is not (necessarily) formative assessment (and vice-versa)," in M. Wilson (ed.), *Towards Coherence between Classroom Assessment and Accountability*, Yearbook of the National Society for the Study of Education, Vol. 103/2, University of Chicago Press, Chicago, IL, pp. 183-188.

Black, P. and **D. Wiliam** (1998), "Assessment and classroom learning", *Assessment in Education: Principles, Policy & Practice*, Vol. 5/1, Routledge, London, pp. 7-74.

Blanchard, M.R. et al. (2010), "Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction", *Science Education*, Vol. 94/4, pp. 577-616.

Blau, D. and **J. Curie** (2006), "Preschool, day care, and afterschool care: Who's minding the kids?", *The National Bureau of Economic Research*, Working Paper No. 10670, Handbook of the Economics of Education, Elsevier, North Holland. www.nber.org/papers/w10670.pdf.

Bloemeke, S. et al. (2012), "Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries", *Teaching and Teacher Education*, Vol. 28/1, pp. 44-55.

Brookhart, S.M. (2007), "Expanding views about formative classroom assessments: A review of the literature", in J.H. McMillan (ed.), *Formative Classroom Assessment*, Teachers College Press, New York, pp. 43-62.

Bryk, A.S. et al. (2010), *Organizing School for Improvement: Lessons from Chicago*, The University of Chicago, IL.

Bryk, A.S. and **M.E. Driscoll** (1988), *The High School as Community: Contextual Influences and Consequences for Students and Teachers*, National Center on Effective Secondary Schools, University of Wisconsin, Madison, WI.

Buccheri, G., N.A. Gürber and **C. Brühwiler** (2011), "The impact of gender on interest in science topics and the choice of scientific and technical vocations", *International Journal of Science Education*, Vol. 33/1, pp. 159-178.

Buczynski, S. and **B. Hansen** (2010), "Impact of professional development on teacher practice: uncovering connections", *Teaching and Teacher Education*, Vol. 26/3, pp. 599-607.

Chaudhury, N. et al. (2006), "Missing in action: Teacher and health worker absence in developing countries", *The Journal of Economic Perspectives,* Vol. 20/1, pp. 91-A4.

Coll, R.K., C. Dahsah and **C. Faikhamta** (2010), "The influence of educational context on science learning: A cross-national analysis of PISA", *Research in Science and Technological Education*, Vol. 28/1, pp. 3-24.

Conley, A.M. et al. (2004), "Changes in epistemological beliefs in elementary science students: Epistemological development and its impact on cognition in academic domains", *Contemporary Educational Psychology*, Vol. 29/2, pp. 186-204.

Creemers, B.P.M. and **G.J. Reezigt** (1997), "School effectiveness and school improvement: Sustaining links", *School Effectiveness and School Improvement*, Vol. 8/4, pp. 396-429.

Creemers, B.P.M. and **L. Kyriakides** (2008), *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*, Routledge, London.

Cunha, F. et al. (2006), "Interpreting the evidence on life cycle skill formation", in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education,* Elsevier, North Holland, pp. 697-812.

Cuevas, P. et al. (2005), "Improving science inquiry with elementary students of diverse backgrounds", *Journal of Research in Science Teaching*, Vol. 42/3, pp. 337-357.

Darling-Hammond, L. (1984), *Beyond the Commission Reports: The Coming Crisis in Teaching*, Rand Corporation, Santa Monica, CA.

Desimone, L.M. et al. (2002), "Effects of professional development on teachers' instruction: Results from a three-year longitudinal study", *Educational Evaluation and Policy Analysis*, Vol. 24/2, pp. 81-112.

Diener, E. and **T. William** (2012), "National accounts of well-being", in K.C. Land, J. Sirgy and A. Michalos (eds), *Handbook of Social Indicators and Quality of Life Research*, Springer, New York.

Dincer, M.A. and **G. Uysal** (2010), "The determinants of student achievement in Turkey", *International Journal of Educational Development*, Vol. 30/6, pp. 592-598.

Drechsel, B., C. Carstensen and **M. Prenzel** (2011), "The role of content and context in PISA interest scales: A study of the embedded interest items in the PISA 2006 science assessment", *International Journal of Science Education*, Vol. 33/1, pp. 73-95.

Eccles, J. S. (1994), "Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices", *Psychology of Women Quarterly*, Vol. 18/4, pp. 585–609.

Eccles, J.S. and **A. Wigfield** (1995), "In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs", *Personality and Social Psychology Bulletin*, Vol. 21/3, pp. 215-225.

Ertesvag, S.K. and **E. Roland** (2015), "Professional cultures and rates of bullying", *School Effectiveness and School Improvement*, Vol. 26/2, pp. 195-214.

European Commission (ed.) (2006), *Science Education Now: A Renewed Pedagogy for the Future of Europe*, European Commission Directorate-General for Research, Brussels, http://ec.europa.eu/research/science-society/document_library/pdf_06/report-rocard-on-science-education_en.pdf (accessed 10 April 2012).

European Commission (ed.) (2004), *Europe Needs More Scientists!*, European Commission, Directorate-General for Research, High Level Group on Human Resources for Science and Technology in Europe, Brussels, http://ec.europa.eu/research/conferences/2004/sciprof/pdf/final_en.pdf (accessed 10 April 2012).

Faubert, V. (2009), "School evaluation: Current practices in OECD countries and a literature review", *OECD Education Working Papers*, No. 42, OECD Publishing, Paris, http://dx.doi.org/10.1787/218816547156.

Fensham, P.J. (2009), "Real world contexts in PISA science: Implications for context-based science education", *Journal of Research in Science Teaching,* Vol. 46/8, pp. 884-896.

Fivush, R. and **N.R. Hamond** (1990), *Knowing and Remembering in Young Children*, University Press, New York, Cambridge.

Fleener, M.J. (1996), "Scientific world building on the edge of chaos: High school students' beliefs about mathematics and science", *School Science and Mathematics*, Vol. 96/6, pp. 312320.

Furtak, E.M. et al. (2012), "Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis", *Review of Educational Research*, Vol. 82/3, pp. 300-329.

Gándara, P. et al. (2003), "English learners in California schools: Unequal resources, unequal outcomes", *Education Policy Analysis Archives*, Vol. 11/36, http://dx.doi.org/10.14507/epaa.v11n36.2003.

Ghuman, S. and **C. Lloyd** (2010), "Teacher absence as a factor in gender inequalities in access to primary schooling in rural Pakistan", *Comparative Education Review,* Vol. 54/4, pp. 539-554.

Gillies, J. and **J.J. Quijada** (2008), "Opportunity to learn: A high impact strategy for improving educational outcomes in developing countries", *Working Paper*, Washington, DC.

Gottfried, A.E. et al. (2009), "A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation", *Journal of Educational Psychology*, Vol. 101/3, pp. 729-739.

Guskey, T.R. (2007), "Multiple sources of evidence: An analysis of stakeholders' perceptions of various indicators of student learning", *Educational Measurement: Issues and Practice*, Vol. 26/1, pp. 19-27.

Gustafsson, J.E. (2008), "Effects of international comparative studies on educational quality on the quality of educational research", *European Educational Research Journal*, Vol. 7/1, pp. 1-17.

Hanushek, E.A. and **L. Wößmann** (2011), "The economics of international differences in educational achievement", in E.A. Hanushek, S. Machin and L. Wößmann (eds.), *Handbook of the Economics of Education,* North Holland, Amsterdam, Vol. 3, pp. 89-200.

Harlen, W. (2007), *Assessment of learning*, Sage, London.

Harlen, W. and **R. Deakin Crick** (2002), *A Systematic Review of the Impact of Summative Assessment and Tests on Students' Motivation for Learning*, review conducted by the Assessment and Learning Research Synthesis Group, EPPI Center, London.

Hattie, J.A.C. (2009), *Visible Learning: A Synthesis of over 800 Meta-analyses Relating to Achievement*, Routledge, London and New York.

Hattie, J. and **H. Timperley** (2007), "The power of feedback", *Review of Educational Research*, Vol. 77/1, pp. 81-112.

Heckman, J. (2006), "Skill formation and the economics of investing in disadvantaged children", *Science*, Vol. 312/5782, pp. 1900-1902.

Heckman J.J., J. Stixrud and **S. Urzua** (2006), "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior", *Journal of Labor Economics*, Vol 24/3, pp. 411-482.

Hill, H.C., B. Rowan and **D.L. Ball** (2005), "Effects of teachers' mathematical knowledge for teaching on student achievement", *American Educational Research Journal,* Vol. 42/2, pp. 371406.

Ho, E.S.C. (2010), "Family influences on science learning among Hong Kong adolescents: What we learned from PISA", *International Journal of Science and Mathematics Education*, Vol. 8/3, pp. 409-428.

Hofer, B.K. and **P.R. Pintrich** (2002), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Hord, S.M. (1997), *Professional Learning Communities: Communities of Continuous Inquiry and Improvement*, Southwest Educational Development Laboratory, Austin, TX.

Ingersoll, R.M. and **D. Perda** (2010), "Is the supply of mathematics and science teachers sufficient?", *American Educational Research Journal*, Vol. 47/3, pp. 563-694.

Jeanpierre, B., K. Oberhauser and **C. Freeman** (2005), "Characteristics of professional development that effect change in secondary science teachers' classroom practices", *Journal of Research in Science Teaching*, Vol. 42/6, pp. 668-690.

Kahneman, D. et al. (2004), "A survey method for characterizing daily life experience: The day reconstruction method", *Science*, Vol. 306/5702, pp. 1776-1760.

King, D. and **M.R. Stephen** (2012), "Learning science through real-world contexts", in B.J. Fraser, K. Tobin and C. McRobbie (eds.), *Springer International Handbooks of Education,* Second International Handbook of Science Education, Springer, Dordrecht, Vol. 24, pp. 69-79.

King, G. and **J. Wand** (2007), "Comparing incomparable survey responses: New tools for anchoring vignettes", *Political Analysis*, Vol. 15, pp. 46-66.

Kingston, N. and **B. Nash** (2011), "Formative assessment: A meta-analysis and a call for research", *Educational Measurement: Issues and Practice*, Vol. 30/4, pp. 28-37.

Kjærnsli, M. and **S. Lie** (2011), "Students' preference for science careers: International comparisons based on PISA 2006", *International Journal of Science Education*, Vol. 33/1, pp. 121-144.

Klieme, E., C. Pauli and **K. Reusser** (2009), "The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms", in T. Janik and T. Seidel (eds.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*, Waxmann, Münster, pp. 137-160.

Kobarg, M. et al. (2011), *An International Comparison of Science Teaching and Learning: Further Results from PISA 2006*, Waxmann, Münster.

Koeppen, K. et al. (2008), "Current issues in competence modeling and assessment", *Journal of Psychology*, Vol. 216/2, pp. 61-73.

Krapp, A. and M. Prenzel (2011), "Research on interest in science: Theories, methods, and findings", *International Journal of Science Education*, Vol. 33/1, pp. 27-50.

Kubiatko, M. and **K. Vlckova** (2010), "The relationship between ICT use and science knowledge for Czech students: A secondary analysis of PISA 2006", *International Journal of Science and Mathematics Education*, Vol. 8/3, pp. 523-543.

Kunter, M. (2005), "Multiple Ziele im Mathematikunterricht", *Pädagogische Psychologie und Entwicklungspsychologie*, Band 51, Waxmann, Münster.

Lavonen, J. and **S. Laaksonen** (2009), "Context of teaching and learning school science in Finland: Reflections on PISA 2006 results", *Journal of Research in Science Teaching*, Vol. 46/8, pp. 922-944.

Lindqvist, E. and **R. Vestman** (2011), "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment", *American Economic Journal: Applied Economics*, Vol. 3/1, pp. 101-128.

Logan, J. et al. (2011), "Children's attendance rates and quality of teacher-child interactions in at-risk preschool classrooms: Contribution to children's expressive language growth", *Child and Youth Care Forum*, Vol. 40/6, pp. 457-477.

Lomos, C., R.H. Hofman and **R.J. Bosker** (2011), "Professional communities and student achievement: A meta-analysis, *School Effectiveness and School Improvement*, Vol. 22/2, pp. 121-148.

Luu, K. and **J.G. Freeman** (2011), "An analysis of the relationship between information and communication technology (ICT) and scientific literacy in Canada and Australia", *Computers and Education*, Vol. 56/4, pp. 1072-1082.

MacKay, R. (2009), "Remarks on the inefficient use of time in the teaching and learning of mathematics in five secondary schools", in J.H. Meyer and A. van Biljon (eds.), *Proceedings of the 15th Annual Congress of the Association for Mathematics Education of South Africa (AMESA): Mathematical Knowledge for Teaching*, pp. 79-85, www.amesa.org.za/AMESA2009/Volume109.pdf#page=85 (accessed 3 April 2012).

Markowitsch, H.J. and **H. Welzer** (2009), *The Development of Autobiographical Memory*, Psychology Press, London.

Martin, M.O., I.V.S. Mullis and **P. Foy** (2008), *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Eighth and Fourth Grades,* Boston College, Chestnut Hill, MA.

McConney et al. (2011), "Bridging the gap? A comparative, retrospective analysis of science literacy and interest in science for indigenous and non-indigenous Australian students", *International Journal of Science Education*, Vol. 33/14, pp. 2017-2035.

McDowall, S. et al. (*2007*), *Evaluation of the Literacy Professional Development Project*, Ministry of Education, New Zealand.

McMillan, J.H. (2001), "Secondary teachers' classroom assessment and grading practices", *Educational Measurement*, Vol. 20, pp. 20-32.

McMillan, J.H. (ed.) (2007), *Formative Classroom Assessment: Theory into Practice*, Teachers College Columbia University, New York.

Millot, B. and **J. Lane** (2002), "The efficient use of time in education", *Education Economics*, Vol. 10/2, pp. 209-228.

Minner, D.D., A.J. Levy and **J. Century** (2010), "Inquiry-based science instruction: What is it and does it matter?", results from a research synthesis 1984 to 2002, *Journal of Research in Science Teaching*, Vol. 47/4, pp. 474-496.

Nagengast, B. and **H.W. Marsh** (2014), "Motivation and engagement in science around the globe: Testing measurement invariance with multigroup structural equation models across 57 Countries using PISA 2006", in L. Rutkowski, M.v Davier and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment*, CRC Press, Boca Raton, FL, pp. 317-344.

Nevo, D. (ed.) (2002), *School-Based Evaluation: An International Perspective*, JAI Press, Amsterdam.

OECD (2014), *PISA 2012 Technical Report*, OECD, Paris, www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf.

OECD (2013), *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 (2011), *Quality Time for Students: Learning In and Out of School*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264087057-en.

OECD (2010a), *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264091450-en.

OECD (2010b), *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes (Volume II)*, PISA, OECD Publishing, http://dx.doi.org/10.1787/9789264091504-en.

OECD (2010c), *The High Cost of Low Educational Performance: The Long-run Economic Impact of Improving PISA Outcomes*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264077485-en.

OECD (2010d), *PISA 2009 Results: What Makes a School Successful?: Resources, Policies and Practices (Volume IV)*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264091559-en.

OECD (2010e), *PISA 2009 Results: Learning to Learn: Student Engagement, Strategies and Practices (Volume III)*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264083943-en.

OECD (2008), *Encouraging Student Interest in Science and Technology Studies*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264040892-en.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264040014-en.

OECD (2006a), *Where Immigrant Students Succeed : A Comparative Review of Performance and Engagement in PISA 2003*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264023611-en.

OECD (2006b), « Improving learning through formative assessment », dans OECD, *Education Policy Analysis 2006: Focus on Higher Education*, OECD Publishing, Paris, http://dx.doi.org/10.1787/epa-2006-5-en.

OECD (2005), "The definition and selection of key competencies: Executive summary", p. 11, www.oecd.org/dataoecd/47/61/35070367.pdf (accessed 2 March 2012).

OECD (2004), *Learning for Tomorrow's World: First Results from PISA 2003*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264006416-en.

Olsen, R.V. and **S. Lie** (2011), "Profiles of students' interest in science issues around the world: Analysis of data from PISA 2006", *International Journal of Science Education*, Vol. 33/1, pp. 97-120.

Osborne, J. (2012), "The role of argument: Learning how to learn in school science", in B.J. Fraser, K. Tobin and C. McRobbie (eds.), *Second International Handbook of Science Education, Springer International Handbooks of Education*, Springer, Dordrecht, Vol. 24, pp. 933-949.

Paine, L. and **L. Ma** (1993), "Teachers working together: A dialogue on organisational and cultural perspectives of Chinese teachers", *International Journal of Educational Research*, Vol. 19/8, pp. 675-697.

Partnership for 21st Century Skills (2008), *21st Century Skills, Education & Competitiveness: A Resource and Policy Guide*, Partnership for 21st Century Skills, Tucson, AZ.

Patall, E.A., H. Cooper and **A.B. Allen** (2010), "Extending the school day or school year", *Review of Educational Research*, Vol. 80/3, pp. 401-436.

Pellegrino, J., N. Chudowsky and **R. Glaser** (eds.) (2001), *Knowing what Students Know: The Science and Design of Educational Assessment*, National Academy Press, Washington, DC.

Poropat, A.E. (2009), "A meta-analysis of the five factor model of personality and academic performance", *Psychological Bulletin*, Vol. 135/2, pp. 322-338.

Porterfield, S. and **A.E. Winkler** (2007), "Teen time use and parental education: Evidence from the CPS, MTF, and ATUS", *Monthly Labor Review,* May 2007, pp. 37-56.

Portner, H. (ed.) (2005), *Teacher Mentoring and Induction: The State of the Art and Beyond*, Corwin, Thousand Oaks, CA.

Prenzel, M., T. Seidel and **M. Kobarg** (2012), "Science teaching and learning: An international comparative perspective", in B.J. Fraser, K. Tobin and C. McRobbie (eds.), *Second International Handbook of Science Education,* Springer International Handbooks of Education, Springer, Dordrecht, Vol. 24, pp. 667-678.

Purves, A.C. (1987), "I.E.A. an Agenda for the Future", *International Review of Education,* Vol. 33/1, pp. 103-107.

Rakoczy, K., E. Klieme and **C. Pauli** (2008), "Die Bedeutung der wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse und des Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation", *Zeitschrift für Pädagogische Psychologie/ German Journal of Educational Psychology*, Vol. 22/1, pp. 25-35.

Richardson, M., C. Abraham and **R. Bond** (2012), "Psychological correlates of university students' academic performance: A systematic review and meta-analysis", *Psychological Bulletin,* Vol. 138/2, pp. 353-387.

Roberts, B.W. et al. (2007), "The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes", *Perspectives on Psychological Science*, Vol. 2/4, pp. 313-345.

Rosenholtz, S.J. (1989), *Teachers' Workplace: The Social Organization of Schools*, Longman, New York.

Rosenkvist, M.A. (2010), "Using student test results for accountability and improvement: A literature review", *OECD Education Working Papers*, No. 54, OECD Publishing, Paris, http://dx.doi.org/10.1787/5km4htwzbv30-en.

Ryan, R.M. and **E.L. Deci** (2000), "Intrinsic and extrinsic motivations: Classic definitions and new directions", *Contemporary Educational Psychology*, Vol. 25/1, pp. 54-67.

Ryan, K.E., M. Chandler and **M. Samuels** (2007), "What should school-based evaluation look like?", *Studies in Educational Evaluation*, Vol. 33/3-4, pp. 197-212.

Rychen, DS. and **L.H. Salganik** (eds.) (2003), *Defining and Selecting Key Competencies*, Contributions to the Second DeSeCo Symposium, Geneva, Switzerland, www.oecd.org/pisa/35070367.pdf.

Salmon, J. et al. (2005), "Reducing sedentary behaviour and increasing physical activity among 10-year-old children: Overview and process evaluation of the 'Switch-Play' intervention", *Health Promotion International,* Vol. 20/1, pp. 7-17.

Sammons, P. et al. (2008), "Children's cognitive attainment and progress in English primary schools during Key Stage 2: Investigating the potential continuing influences of pre-school education", *Frühpädagogische Förderung in Institutionen: Zeitschrift für Erziehungswissenschaft*, VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 179-198.

Sanders, J.R. and **E.J. Davidson** (2003), "A model for school evaluation", in T. Kellaghan, D.L. Stufflebeam and L.A. Wingate (eds), *International Handbook of Educational Evaluation*, Vol. 9, Kluwer International Handbooks of Education, Springer, Dordrecht, pp. 807-826.

Santiago, P. and **F. Benavides** (2009), *Teacher Evaluation: A Conceptual Framework and Examples of Country Practices*, OECD, Paris, www.oecd.org/edu/school/44568106.pdf.

Scheerens, J. (2002), "School self-evaluation: Origins, definition, approaches, methods and implementation", in D. Nevo (ed.), *School-based Evaluation: An International Perspective*, Emerald Group Publishing Limited, Bingley, UK, pp. 35-69.

Scheerens, J. and **Roel J. Bosker** (1997), "The foundations of educational effectiveness", *International Review of Education*, Vol. 45/1, pp. 113-120, http://dx.doi.org/10.1023/A:1003534107087.

Scheerens, J., C.A.W. Glas and **S. Thomas** (eds.) (2003), *Educational Evaluation, Assessment, and Monitoring: A Systemic Approach*, Contexts of Learning, Swets & Zeitlinger, Lisse, Netherlands and Exton, PA.

Scherff, L. and **C.L. Piazza** (2008), "Why now more than ever, we need to talk about opportunity to learn", *Journal of Adolescent and Adult Literacy*, Vol. 52/4, pp. 343-352.

Schleicher, A. (ed.) (2012), *Preparing Teachers and Developing School Leaders for the 21st Century: Lessons from around the World*, International Summit on the Teaching Profession, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264174559-en.

Schmidt, W.H. and **A. Maier** (2009), "Opportunity to learn", Chapter 44, in G. Sykes, B. Schneider and D.N. Plank (eds.), *Handbook of Education Policy Research*, Routledge, New York, pp. 541559.

Schmidt, W.H. et al. (2001), *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*, Jossey-Bass, San Francisco, CA.

Scriven, M. (2003), "Evaluation theory and metatheory", in T. Kellaghan, D.L. Stufflebeam and L.A. Wingate (eds), *International Handbook of Educational Evaluation*, Vol. 9, Kluwer International Handbooks of Education, Springer, Dordrecht, pp. 15-31.

Seidel, T. and **R.J. Shavelson** (2007), "Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results", *Review of Educational Research,* Vol. 77/4, pp. 454-499.

Shayer, M. and **Adhami, M.** (2007), "Fostering cognitive development through the context of mathematics: Results of the CAME project", *Educational Studies in Mathematics*, Vol. 64/3, pp. 265-291, http://dx.doi.org/10.1007/s10649-006-9037-1.

Shepard, L.A. (2006), "Classroom assessment", in R.L. Brennan (ed.), *Educational Measurement*, Praeger Publishers, Westprot, CT, pp. 623-646.

Shulman, L.S. (1985), "Paradigms and research programs in the study of teaching: A contemporary perspective", in M.C. Wittrock, (ed.), *Handbook of Research on Teaching*, 3rd ed, Macmillan, New York, pp. 3-36.

Shute, V.J. (2008), "Focus on formative feedback", *Review of Educational Research*, Vol. 78/1, pp. 153-189.

Smith, P.K. et al. (2002), "Definitions of bullying: Comparison of terms used, and age and gender differences, in a fourteen-country international comparison", *Child Development*, Vol. 73/4, pp. 1119-1133.

Spurrier, N. et al. (2003), "Socio-economic differentials in the health-related quality of life of Australian children: Results of a national study", *Australian and New Zealand Journal of Public Health*, Vol. 27/1, pp. 27-33.

Steinert, B. et al. (2006), "Lehrerkooperation in der Schule" (Teacher co-operation in schools), *Zeitschrift für Pädagogik,* Vol. 52/2, pp. 185-203.

Stigler, J.W. and **J. Hiebert** (1999), *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*, Free Press, New York.

Stiggins, R.J. (2007), "Conquering the formative assessment frontier", in J.H. McMillan (ed.), *Formative Classroom Assessment: Theory into Practice*, Teachers College Columbia University, New York, pp 8-28.

Stiglitz, J.E., A. Sen and **J.-P. Fitoussi** (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, INSEE, Paris.

Supovitz, J.A. and **H.M. Turner** (2000), "The effects of professional development on science teaching practices and classroom culture", *Journal of Research in Science Teaching*, Vol. 37/9, pp. 963-980.

Sykes, G. (1996), "Reform of and as professional development", *Phi Delta Kappan*, Vol. 77/7, pp. 465-467.

Sykes G., B. Schneider and **D.N. Plank** (eds.) (2009), *Handbook of Education Policy Research,* Routledge, New York.

Sylva, K. et al. (2011a), "Pre-school quality and educational outcomes at age 11: Low quality has little benefit", *Journal of Early Childhood Research*, Vol. 9/2, pp. 109-124.

Sylva, K. et al. (2011b), "Effects of early child-care on cognition, language, and task-related behaviours at 18 months: An English study", *British Journal of Developmental Psychology*, Vol. 29/1, pp. 18-45.

Taylor, J.A., M.A.M. Stuhlsatz and **R.W. Bybee** (2009), "Windows into high-achieving science classrooms", in R.W. Bybee and B. McCrae (eds.), *PISA Science 2006, Implications for Science Teachers and Teaching*, NSTA Press, Arlington, VA, pp. 123-132.

The PROBE Team (1999), *Public Report on Basic Education in India,* Oxford University Press, New Delhi.

Timperley, H. et al. (2007), *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration [BES]*, Wellington, Ministry of Education, New Zealand, www.educationcounts.govt.nz/goto/BES (accessed 15 April 2011).

UNESCO (2006), *Education for all Global Monitoring Report 2007: Strong Foundations*, UNESCO Publishing, Paris.

Vieluf, S. et al. (2012), *Teaching Practices and Pedagogical Innovations: Evidence from TALIS*, TALIS, OECD, Paris, http://dx.doi.org/10.1787/9789264123540-en.

Vieluf, S., J. Lee and **P. Kyllonen** (2009), "The predictive power of variables from the PISA 2003 Student Questionnaire", QEG(0910)5a.doc, paper presented at the QEG meeting, Offenbach, Germany, 19-21 October.

Valenti, J.E. and **D.H. Tracey** (2009), "Full-day, half-day, and no preschool effects on urban children's first-grade reading achievement", *Education and Urban Society*, Vol. 41/6, pp. 695-711.

Wang, C., P. Berry and **S.M. Swearer** (2013), "The critical role of school climate in effective bullying prevention", *Theory into Practice*, Vol. 52/4, pp. 296-302.

Wang, L. et al. (2009), "Assessing teamwork and collaboration in high school students: A multimethod approach", *Canadian Journal of School Psychology*, Vol. 24/2, pp. 108-124.

Wigfield, A., J.S. Eccles and **D. Rodriguez** (1998), "The development of children's motivation in school contexts", *Review of Research in Education*, Vol. 23, p. 73.

Willms, J.D. (2010), "School composition and contextual effects on student outcomes", *Teachers College Record*, Vol. 112/4, pp. 1008-1037.

Willms, J.D. (2006), *Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems*, UNESCO Institute for Statistics, Montreal, Canada.

Wilson, M. (ed.) (2004), *Towards Coherence between Classroom Assessment and Accountability*, National Society for the Study of Education Yearbooks, University of Chicago Press, Chicago, IL.

Wößmann, L. et al. (2009), *School Accountability, Autonomy and Choice around the World*, Edward Elgar Publishing, MA.

Wößmann, L. et al. (2007), "School accountability, autonomy, choice, and the level of student achievement: International evidence from PISA 2003", *OECD Education Working Papers*, No. 13, OECD Publishing, Paris, http://dx.doi.org/10.1787/246402531617.

Annex 6.A

Selected analytical models used in publications on the PISA 2006 data for contexts of science achievement

Publication	Research Question or Model
Nagengast and Marsh (2014)	Cross-cultural measurement invariance for motivation and engagement in science
Drechsel, Carstensen and Prenzel (2011)	Dimensionality of science interest
Olsen and Lie (2011)	Country- and culture specific profiles of interest
Ainley and Ainley (2011a)	Students' enjoyment, learning engagement, and achievement
Ainley and Ainley (2011b)	Knowledge, affect, value, and students' interest in science
Lavonen and Laaksonen (2009)	Learning activities, interest in science, self-efficacy, self-concept, and performance
Fensham (2009)	Gender, task context and science performance
Buccheri, Gruber and Bruhwiler (2011)	Gender specificity in interest and vocational choices
Mc Conney et al. (2011)	Science interests among minority students
Luu and Freeman (2011)	Scientific literacy and ICT-related variables
Kubiatko and Vlckova (2010)	
Ho (2010)	Parental involvement and students' science performance
Basl (2011)	Explaining interest in future science-related careers
Kjaernsli and Lie (2011)	
Willms (2010)	School composition, school and classroom context, and students' literacy skills
Dincer and Uysal (2010)	Effects of school programme types
Coll et al. (2010)	influence of educational context in a western vs. Asian country