Chapter 3

Framework Development in International Large-Scale Assessment Studies

John Ainley and Wolfram Schulz

Abstract At the outset of an international large-scale assessment study, researchers typically develop an assessment framework that defines the outcomes (content, skills, and understandings) that are to be assessed in the study and provide a contextual framework that outlines the scope of contexts believed to be associated with those outcomes. A particular challenge is to establish frameworks for cross-national studies that are valid across a wide range of participating education systems. Some studies have assessment frameworks that reference established curriculum areas. Other studies have assessment frameworks that define outcomes related to learning areas of a cross-curricular nature. A third set of studies have assessment frameworks that put a greater emphasis on what people need to know to successfully participate in society rather than on what is taught in schools. Contextual frameworks, typically embedded in a broader assessment framework, focus on those characteristics that are believed to influence student learning either as teaching and learning processes or as the conditions provided for teaching and learning. Some large-scale international surveys have been entirely concerned with contexts and are based on frameworks that are primarily concerned with contexts. Framework development involves detailed definitions of the outcomes to be investigated in terms of content, cognitive processes, and affective dispositions, as well as the types of items to be used and aspects of assessment design. Frameworks are informed by existing research literature and available information about national and school curricula. The development process is informed by expert advisory groups and country representatives through a series of review cycles. Research questions that guide the study are formulated as part of these frameworks.

Keywords Contextual influences • Framework • Learning outcomes • Measurement • Study design

J. Ainley (✉) • W. Schulz
Australian Council for Educational Research (ACER), Camberwell, Australia
e-mail: john.ainley@acer.org

W. Schulz
e-mail: wolfram.schulz@acer.org

3.1 Introduction

Assessment frameworks define not only what is to be measured but also the how and the why of what is to be measured. The international large-scale assessments (ILSAs) of IEA (and those of other organizations such as the Organisation for Economic Co-operation and Development [OECD]) are all based on such an organizing document. Although assessment frameworks are also a common feature of national assessment surveys, they are especially important for ILSAs because of the need to transcend variations among participating education systems. In many ILSAs, there will be an assessment framework that defines the outcomes (content, skills, and understandings) that are to be assessed in the study, as well as a contextual framework that defines the contexts believed to be associated with those measured outcomes. In some studies these are not distinguished, but both elements are essential to underpin the quality of ILSAs and provide stakeholders, the research community, and a wider audience with a rationale for the study, the definitions and conceptualizations of what is measured and for what reason, and an outline of the study design.

A particular challenge specific to cross-national studies is to establish frameworks that are valid across the range of participating education systems. There is considerable diversity across national contexts in the way curricula are defined and implemented, how systems are structured and governed, and in which form teaching and learning take place within schools. The assessment framework is the main point of reference to understand how international studies define common elements of learning outcomes or contextual factors that should be measured, and how the study is designed to obtain comparable data across countries.

3.2 Assessment Frameworks

Assessment frameworks define the intended learning outcomes from an area that is to be assessed as well as the design of the assessment (Jago 2009). In other words, assessment frameworks reference the structure and content of the learning area, as well as defining the design of the assessment (including the population definitions) and detailing how the assessment is to be implemented. Furthermore, these documents are also of central importance as reference points to assess content and construct validity and, in cases of studies that monitor trends over time, they provide rationales for any innovations or changes of content and how these are integrated into the existing construct(s) (Mendelovits 2017).

An assessment framework details the constructs to be measured in terms of the content, skills, and understandings expected to have been learned by the population at the time of their assessment (e.g., grade 8 students). Pearce et al. (2015, p. 110) argued that an assessment framework provides a “structured conceptual map of the learning outcomes” for a program of study. The National Assessment Governing

Board (NAGB) of the United States Department of Education describes the mathematics assessment framework for the National Assessment of Educational Progress (NAEP) as being “like a blueprint” in that it “lays out the basic design of the assessment by describing the mathematics content that should be tested and the types of assessment questions that should be included” (NAGB 2019, p. 2). The Progress in International Reading Literacy Study (PIRLS; see IEA 2020) conducted by IEA similarly says that its assessment framework is “intended as a blueprint” (Mullis et al. 2009, p. 2).

Assessment frameworks are not the same as curriculum frameworks even though the two are related (Jago 2009; Pearce et al. 2015). Whereas a curriculum framework describes comprehensively what is to be taught in education, an assessment framework defines the construct or constructs being measured by a particular assessment and takes into account what can feasibly be assessed. Consequently, assessment frameworks often define the cognitive skills or processes needed by students to deal with relevant tasks in the respective learning area, and some, such as the International Civic and Citizenship Education Study (ICCS; see IEA 2020), also define affective and behavioral outcomes of relevance in the field.

The recent transition of long-term (cyclical) ILSA studies from paper-based to computer-based assessments has provided opportunities to include new measures of aspects with digital technology that could not be administered on paper. While this is an attractive feature of these new ways of assessment delivery, they may also have implications for the construct(s) measured over time. Together with mode effects, computer-enhanced measurement may have consequences for the monitoring of data across assessment cycles and frameworks have the role of describing the extent to which construct(s) may have broadened and include aspects that were not measured in previous cycles. Furthermore, there may also be other aspects that change between ILSA implementations (such as trends in curricular changes across countries or changes in relevance of particular topics for a learning area under study). The assessment framework serves as a reference point for describing any such changes and providing a rationale for them.

Assessment frameworks also incorporate sets of research questions concerned with the expected variations in achievement among and within countries and covariation with student, school, and system characteristics. Those research questions broadly define the analyses to be conducted and guide the structure of reporting.

3.3 Contextual Frameworks

ILSAs study contexts to aid understanding of variation in achievement measures. The constructs that characterize those contexts are elaborated as variables in contextual frameworks on the basis of relevant research literature. The contextual information necessary to provide the basis for measures of the constructs is also outlined in the framework, which is then used as a guide to for the development of questionnaire material (Lietz 2017; see also Chap. 5). For IEA’s assessments, contextual

information is gathered through student, teacher, and school questionnaires, and, in some cases, from parent questionnaires, as well as from information about education systems provided by national research centers and/or experts.

Contextual variables can be classified in various ways. One approach to classification is in terms of what is being measured; for example, whether the variables are factual (e.g., age), attitudinal (e.g., enjoyment of the area), or behavioral (e.g., frequency of learning activities). Another approach recognizes that the learning of individual students is set in the overlapping contexts of school learning and out-of-school learning. A third approach to classifying contextual variables is based on the multilevel structure inherent in the process of student learning (see Scheerens and Bosker 1997; Schulz et al. 2016). It is possible to broadly distinguish the following levels from a multilevel perspective:

- *Context of the individual*: This level refers to the individual ascribed and developed characteristics of individual respondents.
- *Context of home and peer environments*: This level comprises factors related to the home background and the immediate social out-of-school environment of the student (e.g., peer-group activities).
- *Context of schools and classrooms*: This level comprises factors related to the instruction students receive, teacher characteristics, the school culture, leadership, and the general school environment.
- *Context of the wider community*: This level comprises the wider context within which schools and home environments work. Factors can be found at local, regional, and national levels.

Another distinction between types of contextual factors can be made by grouping contextual variables into antecedents or processes (see Fraillon et al. 2019; Schulz et al. 2016):

- Antecedents are those variables from the past that shape how student learning takes place. These factors are level-specific and may be influenced by antecedents at a higher level. For example, training of teachers may be affected by historical factors and/or policies implemented at the national level. In addition, educational participation may be an antecedent factor in studies beyond the compulsory years of schooling.
- Processes are those variables related to student learning and the acquisition of understandings, competencies, attitudes, and dispositions. They are constrained by antecedents and possibly influenced by variables relating to the higher levels of the multi-level structure.

Antecedents and processes are variables that have potential impact on the outcomes at the level of the individual student. Learning outcomes at the student level can also be viewed as aggregates at higher levels (school, country) where they can affect factors related to learning processes.

3.4 Design and Implementation

Being blueprints for ILSA studies, frameworks also describe assessment designs and implementation strategies. An assessment design will include specification of the balance of numbers of items across content and cognitive domains, the types of items (closed or constructed response) and details of a rotated block design for studies where such a design is to be employed. Since the first cycle of the Trends in International Mathematics and Science Study (TIMSS; see IEA 2020) in 1995, most ILSA studies have employed rotated block designs, in which students are randomly allocated booklets based on different blocks of items, so as to provide a more comprehensive coverage of the learning area than could be assessed with a single set of items. An assessment framework for an ILSA study would typically also indicate how scale scores are to be calculated from item response data to sets of items related to the constructs being measured.

An assessment framework may also include a description of available delivery modes (paper-based or computer-based). The International Computer and Information Literacy Study (ICILS; see IEA 2020) has been computer-based since its inception and requires relevant specification of item types (Fraillon et al. 2019). Computer-based methods have since become part of other ILSA studies. Mixed-mode methods, or changes in mode between cycles, require consideration of whether delivery mode influences student responses and how to evaluate any mode effects (see also Chap. 10).

Implementation strategies described in assessment framework documents include population definition, sample design, and administration procedures. For example, frameworks for IEA study populations would provide definitions in terms of the grade to be assessed (which typically includes either grade 4 or grade 8, or both grade 4 and grade 8), while frameworks for the OECD's Programme for International Student Assessment (PISA) define the population in terms of student age (15 years) and secondary school attendance at the time of testing. It is important that the framework provides explicit information about the population and whether exclusions are permitted (e.g., students with special educational needs). Most ILSA studies allow exclusions that meet specific criteria for up to five percent of the population.

Assessment frameworks also tend to provide outlines of the instrumentation and other design features (such as international options) to participating countries (e.g., including additional sets of questions as part of the questionnaires). It is further customary to provide examples of assessment material, which illustrate the way constructs are measured in the respective ILSA. By giving insights into the ways of measuring these constructs, ILSAs provide stakeholders, researchers, and the wider audience with a better understanding of their operationalization and enable them to assess their validity, in particular in cases where new innovative construct(s) are measured (see, e.g., Fraillon et al. 2019).

3.5 Steps in Framework Development

Framework development is based on, and makes use of, understandings of the research literature about the respective learning area and information about educational practice and policy in that field. These sources provide the basis for a proposal for research and help to define the scope of the framework, which guides the assessment and identifies the contextual influences to be measured along with the assessment. They help to identify the content and skills to be assessed in an ILSA study and the basis for the design and format of the assessment (Jago 2009). The process of reviewing literature and educational practice results in the formulation of research questions that become key elements of the assessment framework. At the broadest level this involves a definition of the main construct(s) an ILSA sets out to measure. For example, based on an extensive literature review, PIRLS defines reading literacy as:

the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment (Mullis and Martin 2015, p. 12).

A broad definition provides the basis for a detailed elaboration of the construct that is embedded in theory. For example, in PIRLS, reading for literary experience and reading to acquire and use information each incorporate four types of comprehension process: (1) retrieving explicitly stated information, (2) making straightforward inferences, (3) interpreting and integrating ideas and information, and (4) evaluating content and textual elements (Mullis and Martin 2015). Research literature, theory, and information about educational practice provide the basis for these elaborations. They also provide a basis for systematically characterizing important aspects of contexts and articulating the research questions for the respective ILSA.

Although the processes of reviewing literature are well established, processes for reviewing educational practice and policy are less clearly set out. ILSA studies frequently make use of reviewing national curricula and curriculum frameworks to augment the insights provided by literature reviews. Reviewing curricula through document analysis is clearer when there is a direct alignment between a discipline or subject and the construct being assessed. This is the case for mathematics and science as assessed in TIMSS (Mullis and Martin 2017). It is also evident when the capability being assessed is a key element of all primary school curricula and part of the language arts learning area as in the assessment of reading in PIRLS (Mullis and Martin 2015). It is less clearly evident when the capability (or assessment construct) being assessed crosses several curriculum areas as is the case in ICCS (Schulz et al. 2016) and ICILS (Fraillon et al. 2019). In studies that set out to measure so-called real-life skills, such as the OECD PISA study or the OECD Programme for the International Assessment of Adult Competencies (PIAAC), the respective frameworks need to provide a theoretical underpinning for the respective conceptualizations of construct(s) that are not referenced to any national curricular frameworks but rather to overarching formulations of expected knowledge and skills based on theoretical

models (OECD 2012, 2019). The increasing heterogeneity among participants in ILSA studies means that the processes to ensure inclusion when defining content and design are crucial to the validity of the assessments.

Structured opportunities for country commentary and expert advice provide important perspectives that contribute to framework development. ILSA studies conducted by the IEA incorporate reviews by national research coordinators (NRCs) so that the framework is applicable to all of the participating countries. NRCs meet in person several times during an ILSA study and communicate outside those meetings on a regular basis. It also essential to seek expert advice during the process of framework development. PIRLS has a reading development group and a questionnaire development group that contribute to framework development (Mullis and Martin 2015). TIMSS has a science and mathematics item review committee and a questionnaire item review committee that contribute to framework development (Mullis and Martin 2017). ICILS and ICCS each have project advisory committees providing expert advice across all aspects of these studies. OECD studies such as TALIS, PISA, and PIAAC have established similar expert groups to provide advice on the development of their respective frameworks.

3.6 Types of Framework

Among the range of ILSA studies there appear to be four main types of framework that reflect the nature of the domains of those studies.

3.6.1 *Curriculum-Referenced Frameworks*

TIMSS and PIRLS, the longest established of the ILSA studies currently conducted by IEA, are closely related to learning areas of school curricula. For example, the TIMSS curriculum model (Mullis and Martin 2017), which is also present in similar forms in other IEA studies, includes three aspects: the intended curriculum, the implemented curriculum, and the attained curriculum. These represent, respectively, the domain-related aspects that students are expected to learn, what is actually done in classrooms, the characteristics of those facilitating these learning opportunities, how it is taught and offered, what students have learned, what they think about learning these subjects, and how they apply their knowledge.

TIMSS is most clearly aligned with mathematics and science at grades 4 and 8. The assessment frameworks are organized around a content and a cognitive dimension. The content domains for mathematics at grade 4 are number (50%), measurement and geometry (30%), and data (20%), and the cognitive domains are knowing (40%), applying (40%), and reasoning (20%) (Mullis and Martin 2017). At grade 8, the content domains are number (30%), algebra (30%), geometry (20%), and data and probability (20%), and the cognitive domains are knowing (35%), applying (40%),

and reasoning (25%) (Mullis and Martin 2017). In summary there is less emphasis on number at grade 8 than there is at grade 4. At grade 8, there is a little more emphasis on reasoning, and a little less emphasis on knowing than at grade 4.

The content domains for TIMSS science at grade 4 are life science (45%), physical science (35%), and earth science (20%), and the cognitive domains are knowing (40%), applying (40%), and reasoning (20%) (Mullis and Martin 2017). At grade 8, the content domains are biology (35%), chemistry (20%), physics (25%), and earth science (20%) and the cognitive domains are knowing (35%), applying (35%), and reasoning (30%). In brief, at grade 8 there is a little less emphasis on biology (or life science) and a little more emphasis on physics and chemistry (physical science), as well as greater emphasis on reasoning and less on knowing than at grade 4. The context framework for TIMSS is organized around five areas: student attitudes to learning, classroom contexts, school contexts, home contexts, and community or national policies (Mullis and Martin 2017).

PIRLS is only designed as a grade 4 assessment and defines its content as focused on reading for literary experience and reading to acquire and use information, with each incorporating four types of comprehension process: retrieving explicitly stated information, making straightforward inferences, interpreting and integrating ideas and information, and evaluating content and textual elements (Mullis and Martin 2015). The framework specifies that 50% of the assessment is concerned with reading for literary experience and 50% is concerned with reading to acquire and use information (Mullis and Martin 2015). The framework identifies four comprehension processes: retrieving explicitly stated information (20%), making straightforward inferences (30%), interpreting and integrating ideas and information (30%), and evaluating content and textual elements (20%) (Mullis and Martin 2015). The PIRLS 2016 framework observes that, in the literature on reading, there has been a shift in emphasis from fluency and basic comprehension to demonstrating the ability to apply what is read to new contexts and that this has been reflected in the framework.

3.6.2 Frameworks for Measuring Outcomes in Cross-Curricular Learning Areas

While, in spite of cross-national diversity, there is considerable common ground across different national curricula with regard to subject areas such as mathematics or science, there are also learning areas that are cross-curricular and/or embedded in different subjects or subject areas. Two such learning areas that are increasingly regarded as highly relevant for education across a wide range of societies are students' competencies related to civic and citizenship education and their information and communication technology (ICT) related skills. The IEA has recognized the importance of investigating these two areas through the establishment of two continuous studies: ICCS (with its first two completed cycles in 2009 and 2016, and an upcoming

cycle in 2022) and ICILS (with its first two completed cycles in 2013 and 2018, and an upcoming cycle in 2023).

Given the diversity of approaches to these learning areas across national contexts (Ainley et al. 2013; European Commission/EACEA/Eurydice 2017), it was necessary to develop frameworks for each that were appropriate, relevant, and accepted by participating countries, while at the same time recognizing the existence of a wide range of different approaches to teaching and learning of relevant outcomes. To ensure the validity of frameworks that needed to outline common ground for measurement for these studies, it was necessary to implement an iterative process of reviews and revisions with input and feedback from a range of international experts and national centers in each participating country.

ICCS 2009 was established as a baseline study of ongoing comparative research of civic and citizenship education; it built on previous IEA studies of civic and citizenship education conducted in 1971 as part of the six-subject study (Torney et al. 1975) and in 1999 (the Civic Education Study [CIVED]; see IEA 2020; Torney-Purta et al. 2001). To measure cognitive civic learning outcomes, the ICCS assessment frameworks (Schulz et al. 2008, 2016) articulated what civic knowledge and understanding comprised in terms of two cognitive domains that distinguish knowledge related to both concrete and abstract concepts (knowing) from cognitive process required to reach broader conclusions (reasoning and applying). Furthermore, it distinguished four content domains: civic society and systems, civic principles, civic participation, and civic identities. For both content and cognitive domains, the definition focused on issues that could be generalized across societies and excluded nationally specific issues, such as those related to the particular form of government in a country (for details on the measurement of understanding of civics and citizenship, see Schulz et al. 2013).

For a study of civic and citizenship education it is also paramount to appropriately consider affective-behavioral outcomes. ICCS considers attitudes and engagement-related indicators as part of the learning outcomes that need to be measured through data from its student questionnaires. To this end, the assessment framework of this study defines two affective-behavioral domains (attitudes and engagement) that comprise different indicators of relevant learning outcomes. Furthermore, the contextual framework describes a wide range of factors at different levels (the individual, home, and peer context, the school and classroom context, and the context of the wider community ranging from local communities to supra-national contexts) distinguishing antecedents and process-related variables (Schulz et al. 2016).

While civic and citizenship education is a long-established learning area in a wide range of education systems, education for learning about ICT or digital technologies is a more recent development that followed their emergence as important for people's daily lives. Across many education systems the area has been acknowledged as of importance for young people's education, although there has been a diversity of approaches (see, e.g., Bocconi et al. 2016). ICT-related learning is often envisaged as a transversal or cross-curricular skill and ICT subjects are not consistently offered across countries (Ainley et al. 2016).

To capture to the essence of a number of previous conceptualizations, ICILS 2009 defined computer and information literacy (CIL) as “an individual’s ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society” (Fraillon et al. 2013, p. 17). While continuing the measurement of CIL in its second cycle, ICILS 2018 also included an optional assessment of computational thinking (CT), defined as referring to “an individual’s ability to recognize aspects of real-world problems which are appropriate for computational formulation and to evaluate and develop algorithmic solutions to those problems so that the solutions could be operationalized with a computer” (Fraillon et al. 2019, p. 27).

Both CIL and CT are described in separate sections of the assessment framework, built on previous conceptualizations, and developed in close cooperation with experts and representatives of participating countries. ICILS also studied the use of digital technologies as a means of continuing earlier IEA research from the 1990s and early 2000s, such as that of the Second International Technology in Education Study (SITES; see IEA 2020).

These two examples of learning areas that are transversal and cross-curricular illustrate that a consistent curricular-referenced approach to the development of instruments to measure outcomes is often difficult to implement. Rather, for both ICCS and ICILS the approach is to develop definitions of constructs for measurement that are regarded as relevant to the diversity of approaches across national curricula. In this respect, the development of assessment frameworks that clearly describe the scope of measurement is of particular importance in terms of the validity of the study results.

3.6.3 Frameworks for Measuring Real-Life Skills

OECD studies of educational outcomes and learning tend to define ranges of previously defined knowledge skills that are viewed as important for citizens instead of referencing these skills to existing curricula. OECD’s PISA, which has been conducted every three years since 2000, is designed to assess the extent to which 15-year-old students near the end of their compulsory education “have acquired the knowledge and skills that are essential for full participation in modern societies” (OECD 2019, p. 11). While the study routinely assesses reading, mathematics, and science literacy in each cycle, in particular cycles it also has assessed additional domains such as problem solving, digital reading, or financial literacy. In each cycle, one of the core domains is assessed with more extensive coverage as the major domain, while the two others are measured with less item material.

In its approach to the assessment of educational achievement, PISA emphasizes its policy-orientation that links outcomes to characteristics of students, schools, and education systems, its concept of “literacy” referring to the capacity of 15-year-old students to apply knowledge to different real-world situations, its relevance to life-long learning concepts, its regular assessments and breadth of coverage. While each

domain is defined and described in separate frameworks, the contextual aspects for measurement in each PISA cycle are outlined in a questionnaire framework that focuses on variables that are particularly relevant for the major domain (such as reading literacy in PISA 2018).

The assessment of adults' knowledge and skills in OECD's PIAAC study is administered every ten years and aims at the measurement of key cognitive and workplace skills that are regarded as necessary for participation in society and prosperity at the system level (OECD 2012). As with PISA, the framework defines the competences (without reference to national curricula) and measures the extent to which adults demonstrate them. In its last cycle in 2011/2012, PIAAC set out to measure adult skills in literacy, numeracy and problem solving in technology-rich environments across 25 countries. Again, its assessment framework provides an important reference point for understanding the study's results, as it illustrates the scope of skills that are assessed in terms of what are considered relevant adult skills for participation in modern societies.

While curriculum-referenced studies can be judged in terms of their coverage of what education systems have defined as aims for learning, the OECD approach to assessing achievement sets out overarching (international) learning goals that are defined by looking at what "should" be expected from an "output" perspective. In this respect, the frameworks for PISA and PIAAC were both shaped by the OECD Definition and Selection of Competencies (DeSeCo) project (Rychen and Salganik 2003). Both PISA and PIAAC explicitly define ranges of knowledge and skills that are deemed essential for young people and adults instead of referencing existing curricula. The theoretical bases for their respective frameworks need to be more extensively elaborated than is the case for curricula-referenced frameworks and are of key importance for an understanding of their results.

3.6.4 Frameworks for Measuring Contexts

The OECD Teaching and Learning International Survey (TALIS) and IEA's SITES Module 2 (SITES-M2) are studies of contexts that have not included assessment data.

TALIS is an ongoing large-scale survey of teachers, school leaders, and their learning environments. TALIS was first administered in 2008, and then again in 2013 and 2018. It has been administered in lower secondary schools (ISCED level 2 according to the international ISCED classification, see UNESCO Institute of Statistics 2012) with options to be administered in primary (ISCED level 1) and upper secondary (ISCED level 3) schools and a further option for the survey to be administered in PISA sampled schools. The TALIS 2018 framework built on the cycles in 2008 and 2013. It was developed with advice from a questionnaire expert group (through a series of virtual and personal meetings), national project managers and the OECD secretariat (Ainley and Carstens 2018). The development focus was on effective instructional and institutional conditions that enhance student learning and how these vary within and across countries and over time.

The TALIS 2018 framework addressed enduring themes and priorities related to professional characteristics and pedagogical practices at the institutional and individual levels: teachers' educational backgrounds and initial preparation; their professional development, instructional and professional practices; self-efficacy and job satisfaction; and issues of school leadership, feedback systems, and school climate. It also addressed emerging interests related to innovation and teaching in diverse environments and settings.

SITES-M2 was a qualitative study of innovative pedagogical practices using ICT (Kozma 2003). It aimed to identify and describe pedagogical innovations that were considered valuable by each country and identify factors contributing to the successful use of innovative technology-based pedagogical practices. The framework specified procedures for identifying innovative practices in teaching classrooms in primary, lower-secondary, and upper-secondary schools and methods to be used to collect and analyze data. National research teams in each of the participating countries applied these common case study methods to collect and analyze data on the pedagogical practices of teachers and learners, the role of ICT in these practices, and the contextual factors supporting and influencing them. Data were collected from multiple sources for each case, including questionnaires for school principals and technology coordinators, individual or group interviews, classroom observations, and supporting materials (such as teacher lesson plans).

3.7 Conclusions

Most research studies are based on frameworks that link them to the extant literature in the field, define research questions, and articulate methods. However, frameworks are especially important for ILSA studies because of the need to ensure validity across a great diversity of national contexts and education systems. Frameworks for ILSA studies need to be explicit about the constructs being measured and the ways in which they are measured. An assessment framework should be the main point of reference to understand how common elements of learning are defined and measured, and how comparable data across countries are to be generated and analyzed.

Consistency of definition and measurement is already a challenge, even for achievement in fields such as mathematics and science, but it is more of a challenge in learning areas that are context-dependent, such as reading, and/or of a rather transversal, cross-curricular nature and not consistently articulated in curriculum documents, such as civic and citizenship education and ICT-related skills. The importance of assessment frameworks for providing reference points in terms of construct validity is also highlighted in cases where ILSAs need to document content- and method-related changes in terms of the definition of what and how learning outcomes are measured (such as with the transition to computer-based delivery of assessments or the adaptation of new content areas as a consequence of societal developments that affect the respective learning area).

References

- Ainley, J., & Carstens, R. (2018). *Teaching and learning international survey (TALIS) 2018 conceptual framework*. OECD Education Working Papers, No. 187. Paris, France: OECD Publishing. <http://dx.doi.org/10.1787/799337c2-en>.
- Ainley, J., Fraillon, J., Schulz, W., & Gebhardt, E. (2016). Conceptualizing and measuring computer and information literacy in cross-national contexts. *Applied Measurement in Education*, 29(4), 291–309. <https://doi.org/10.1080/08957347.2016.1209205>.
- Ainley, J., Schulz, W., & Friedman, T. (Eds.). (2013). *ICCS 2009 Encyclopedia. Approaches to civic and citizenship education around the world*. Amsterdam, The Netherlands: IEA. <https://www.iea.nl/publications/encyclopedias/iccs-2009-encyclopedia>.
- Bocconi, S., Chiocciariello, A., Dettori, G., Ferrari, A., Engelhardt, L., Kampylis, P., et al. (2016). Developing computational thinking: Approaches and orientations in K-12 education. In G. Veletianos (Ed.), *EdMedia. World Conference on Educational Media and Technology 2016, June 28–30, 2016, Vancouver, British Columbia* (pp. 13–18). Waynesville, NC: Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/j/EDMEDIA/v/2016/n/1>.
- European Commission/EACEA/Eurydice. (2017). *Citizenship education at school in Europe—2017*. Eurydice Report. Luxembourg: Publications Office of the European Union. https://eacea.ec.europa.eu/national-policies/eurydice/content/citizenship-education-school-europe-%E2%80%9393-2017_en.
- Fraillon, J., Ainley, J., Schulz, W., Duckworth, D., & Friedman, T. (2019). *IEA international computer and information literacy study 2018 assessment framework*. Cham, Switzerland: Springer. <https://www.iea.nl/publications/assessment-framework/icils-2018-assessment-framework>.
- Fraillon, J., Schulz, W., & Ainley, J. (2013). *International computer and information literacy study 2013: Assessment framework*. Amsterdam, the Netherlands: IEA. <https://www.iea.nl/publications/assessment-framework/international-computer-and-information-literacy-study-2013>.
- IEA. (2020). *IEA studies [webpage]*. Amsterdam, the Netherlands: IEA. <https://www.iea.nl/index.php/studies/ieastudies>.
- Jago, C. (2009). *A history of NAEP assessment frameworks*. Paper commissioned for the 20th anniversary of the National Assessment Governing Board 1988–2008. Washington, DC: National Assessment Governing Board, US Department of Education. www.nagb.gov/focus-areas/reports/history-naep-assessment-frameworks.html.
- Kozma, R. (Ed.). (2003). *Technology, innovation, and educational change, a global perspective. A report of the Second Information Technology in Education Study Module 2*. Eugene, OR: International Society for Technology in Education (ISTE).
- Lietz, P. (2017). Design, development and implementation of contextual questionnaires in large-scale assessments. In P. Lietz, J. Cresswell, K. Rust, & R. Adams (Eds.), *Implementation of large-scale education assessments* (pp. 92–136). Chichester, UK: Wiley.
- Mendelovits, J. (2017). Test development. In P. Lietz, J. Cresswell, K. Rust, & R. Adams (Eds.), *Implementation of large-scale education assessments* (pp. 63–91). Chichester, UK: Wiley.
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://www.iea.nl/index.php/publications/assessment-framework/pirls-2016-assessment-framework>.
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://www.iea.nl/index.php/publications/assessment-framework/timss-2019-assessment-frameworks>.
- Mullis, I., Martin, M., Kennedy, A., Trong, K., & Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <https://www.iea.nl/index.php/publications/assessment-framework/pirls-2011-assessment-framework>.

- NAGB. (2019). *Mathematics framework for the 2019 National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board, U.S. Department of Education. www.nagb.gov/content/nagb/assets/documents/publications/frameworks/mathematics/2019-math-framework.pdf.
- OECD. (2012). *Literacy, numeracy and problem solving in technology-rich environments: Framework for the OECD survey of adult skills*. Paris, France: OECD Publishing. <https://doi.org/10.1787/9789264128859-en>.
- OECD. (2019). *PISA 2018 assessment and analytical framework*. Paris, France: OECD Publishing. <https://doi.org/10.1787/b25efab8-en>.
- Pearce, J., Edwards, J., Fraillon, J., Coates, H., Canny, B., & Wilkinson, D. (2015). The rationale for and use of assessment frameworks: improving assessment and reporting quality in medical education. *Perspectives on Medical Education*, 4, 110–118. <https://link.springer.com/article/10.1007/s40037-015-0182-z>.
- Rychen, D. S., & Salganik, L. H. (Eds.). (2003). *Key competencies for a successful life and a well-functioning society*. Göttingen, the Netherlands: Hogrefe and Huber.
- Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford, UK: Pergamon.
- Schulz, W., Ainley, J., Fraillon, J., Losito, B., & Agrusti, G. (2016). *IEA international civic and citizenship education study 2016 assessment framework*. Cham, Switzerland: Springer. <https://www.iea.nl/publications/assessment-framework/iea-international-civic-and-citizenship-education-study-2016>.
- Schulz, W., Fraillon, J., & Ainley, J. (2013). Measuring young people's understanding of civics and citizenship in a cross-national study. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 33(3), 327–349.
- Schulz, W., Fraillon, J., Ainley, J., Losito, B., & Kerr, D. (2008). *International civic and citizenship education study 2009: Assessment framework*. Amsterdam, The Netherlands: IEA. <https://www.iea.nl/publications/assessment-framework/international-civic-and-citizenship-education-study-2009>.
- Torney, J., Oppenheim, A. N., & Farnen, R. F. (1975). *Civic education in ten countries: An empirical study*. New York, NY: Wiley.
- Torney-Purta, J., Lehmann, R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries*. Amsterdam, the Netherlands: IEA. <https://www.iea.nl/publications/study-reports/international-reports-iea-studies/citizenship-and-education-twenty-eight>.
- UNESCO Institute for Statistics. (2012). *International standard classification of education. ISCED 2011*. Montreal, Canada: UNESCO Institute for Statistics. <http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.