SDG 4.2.1: Connecting *early learning* to the UIS Reporting Scales

Purpose of this paper

This paper describes a way forward for SDG 4.2.1 interim reporting, using the UNESCO Institute for Statistics (UIS) Reporting Scales (UIS RS). Its focus is twofold:

- To leverage existing tools, including the UIS RS and *Data Alignment for SDG 4.1.1 reporting* (ACER, 2017) to support greater consistency in interim reporting against 4.2.1, and to provide a structure to support capacity development within countries to strengthen the quality of the data used for 4.2.1 reporting.

- To establish a framework to describe a continuum of early learning and to support the description of what is meant be *developmentally on track*.

The consensus for SDG 4.1.1 reporting is to "be as pragmatic as possible while being as rigorous as possible" and to "build on existing work and what is available" (GAML Cross National Assessments Technical Meeting, Hamburg, 2017). This is consistent with a fit-for-purpose approach to SDG 4 reporting that prioritises the inclusion of as many countries as possible, and focusing on iterative capacity building to improve the quality and consistency of data used for reporting. In 4.1.1 reporting, this includes acknowledging that education systems make independent decisions about what learning means and how it is assessed. The agreed approach must therefore be flexible enough to accommodate this variety of assessment programs, because no single measure will ever achieve coverage of all countries. The UIS reporting scales (UIS RS) are a key component of SDG 4.1.1, in that they enable different assessments to be aligned with a common scale. This enhances consistency in SDG 4.1.1 reporting, in the context of diverse assessment programs and results.

ACER-GEM's proposal to establish a framework for 4.2.1 interim reporting that is linked to the UIS RS is based upon a similar approach. It leverages the existing work on the UIS RS to describe a continuum of holistic learning outcomes that is based in theory and empirical studies. From this linkage, existing measurements of early learning can be *aligned* (conceptually or empirically) to support countries to use their available data to report against 4.2.1. Aligning measures with the UIS RS will provide a description of the distribution of skills and capabilities observed in children and therefore provide evidence for a criterion-based definition of *developmentally on track* for learning. Such an approach also allows for a mapping of definitions of proficiencies described in the range of assessment tools used for interim reporting.

The approach described in this paper is consistent with the remit of the GAML 4.2 Task Force Terms of Reference which focuses specifically on *learning* within 4.2.1, and how connections to SDG 4.1 methodologies can support interim reporting.

Indicator 4.2.1: Measuring early learning

The UNESCO Institute for Statistics (UIS) is responsible for defining and measuring globally-comparable indicators of Sustainable Development Goal (SDGs) 4. The Global Alliance to Monitor Learning (GAML) brings together stakeholders to build consensus on how to produce globally-comparable data on *learning* for each target within SDG 4. The Terms of Reference for GAML Task Force 4.2 require the group to consider

- *the comparability of measurement across countries;*
- *how best to define the proficiency level, or in the case of 4.2, defining "developmentally on track,"; and*
- *how frequently data should be collected to effectively track changes in outcomes due to interventions, and to guide policy*

UNICEF is the custodian agency identified by the UN Statistical Commission for 4.2.1, and has commissioned an Expert Advisory Panel and Inter-Agency Expert Group on Early Childhood Development (IAEG-ECD) to revise and expand the Early Childhood Development Index (ECDI) with the goal of using it as the indicator to monitor progress on 4.2.1.

One of Task Force 4.2's key tasks is to propose an interim measurement strategy for 4.2.1 until the ECDI revision is complete. This includes options for integration of existing data to address the need for learning indicators. In addition, the Task Force is charged with helping link methodological developments relating to learning for pre-school ages to those for early primary grades through target 4.1.1. The approach proposed in this paper supports all of these objectives.

SDG indicator 4.2.1, the '*Proportion of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being, by sex*', has a scope that goes beyond the learning domain. The focus of Task Force 4.2 is on learning, so other domains are not explicitly addressed in this paper, although options may be considered for how this approach may apply to them.

Key Issues:

- **The revised ECDI is not immediately available and an interim strategy for reporting is required.**
- **A framework to describe early learning is required to support the definition of *on track* and to align the mix of instruments used for interim reporting against 4.2.1.**
- **In the medium term, it is unlikely that the ECDI will cover all countries. There is a need to consider how countries that do not use the ECDI can report against 4.2.1.**
- **Leveraging solutions designed for 4.1.1 will support efficiency and consistency in global reporting, and promote greater continuity between early years and school-age learning.**

UIS reporting scales

The UIS RS are common scales that describe levels of proficiency in reading and mathematics, from the early to secondary years of school education. They are designed to support reporting against Indicator 4.1.1, as well as broader capacity building in learning assessments.

The UIS RS are being developed by UNESCO Institute for Statistics (UIS) with its technical partner, the Australian Council for Educational Research Centre for Global Education Monitoring (ACER-GEM). Draft scales for reading and mathematics have been created, to be validated and refined in 2018 onwards. The draft scales can be viewed in the prototype Learning Progression Explorer (https://learning-progress-explorer.acer.org).

The UIS RS are designed to be used in 4.1.1 reporting in three possible ways:

1. Countries may use an assessment program that has been empirically equated with the UIS RS

2. Countries may use an item pool to empirically link other or new assessments to the UIS RS

3. Countries may undergo a qualitative process to examine the conceptual alignment of an assessment with the UIS RS.

How can the UIS Reporting Scales assist Task Force 4.2

The UIS RS approach has four main benefits in relation to Indicator 4.2.1:

1. Balancing consistency and flexibility

The development of the UIS RS is driven by the need for indicators of learning to be consistent internationally, but flexible enough to allow each country to pursue assessment programs that fit their unique context. This same principle—balancing consistency and flexibility—is equally important in the measurement of early learning, recognises that countries may measure learning in different ways. While greater consistency in the use of a single measures (such as ECDI in the case of 4.2.1) may be a desirable long-term goal, the UIS RS recognises that diversity in learning assessment is likely to remain a challenge for SDG reporting over time.

2. Supporting a shared understanding of early learning

There is wide variation across (and within) countries in how early learning is conceptualised. Extending the UIS RS into the early years will support a common global understanding, owned by all members of the GAML network, of what early learning entails. While different understandings of early learning will continue to exist in different contexts, it will ensure that international discussion about measurement of learning against Indicator 4.2.1 uses a common language.

3. Supporting iterative improvement

The development of the UIS RS is designed to be an iterative, multi-year program, resulting in increasing consistency in international indicators of learning over time. This iterative approach begins with the conceptual development of the draft scales, which will then be empirically validated and further refined. Each time the scales are applied in the field, more data is gained to refine the descriptions of learning at each level, and to identify assessment items that can be used to test them. This same iterative approach has particular value for early learning, as a field in which assessment strategies are rapidly evolving and improving. The UIS RS will enable measurement against 4.2.1 to keep pace with changes in the field.

4. Supporting connections between early learning and school education

Applying the UIS RS to measures of early learning for Indicator 4.2.1 will help to connect early learning with school-age education. This will have particular value

in helping education systems to monitor how well early learning experiences are preparing children for subsequent learning, and to identify whether differences between student cohorts at school in fact begin in the early years.

Extending the UIS RS to early learning

The draft UIS RS is currently intended to describe down to the earliest level of proficiency that is likely to arise in measures of learning in reading and mathematics. In order to adapt this to the holistic concept of learning adopted in 4.2.1, there are two major considerations. The first is a consideration of whether the UIS RS reports "low" enough to capture the range of capabilities exhibited by children *under five*. The second is whether the descriptions of reading and mathematics are broad enough to capture the associated general skills (for example, early cognitive capabilities) that precede the emergence of the specific academic skills described in the UIS RS. Each of these issues is addressed below.

Describing capabilities *under five*

Level 1 through 3 of the UIS RS represent foundational skills in reading and mathematics, and many of the descriptions of capabilities at these levels are consistent with pre-reading and pre-mathematics. These descriptions likely represent capabilities relevant to children in the third – fifth year of life.

In the UIS RS for reading, for example, within the comprehension strand, the descriptions clearly include the pre-reading skills of listening and attending to aural stories, remembering key pieces of information (and using non-verbal pointing or expressive language to respond), using judgements to describe things that are liked, or extrapolating to describe what may happen. An example is provided below of a skill illustration for the comprehension strand at Level 1 and 2 respectively:

- Students listen to the story (4 panel picture story) and then answer aural questions. For example, "Why does (the character) run home?".

- Students listen to the story (approx.10 sentences with illustrations) and make a simple inference, that (the character) does not like his job because he was so bored that he fell asleep.

In the constrained skills strand, other pre-reading skills are in evidence, including book and text knowledge, letter recognition, and phonemic awareness.

Similarly, pre-mathematics skills are evidenced in the early levels of the UIS RS for mathematics. For example, in the measurement strand, skills where children exhibit simple strategies to make comparisons and count can be seen in Level 1 and 2 respectively:

- This task requires students to understand information presented verbally in a single, short sentence, including the term 'smallest', and to interpret images in relation to the relevant variable; devise and apply a simple comparison strategy to identify the 'smallest'; and select their response by choosing one out of four presented shapes.

- This task requires students to understand information presented verbally in a single short sentence, including the relational term 'most'; interpret quantities of a familiar real-world object represented graphically in different arrangements; find a single step strategy to identify the group with 'the most' (eg, count and compare); and select their response from four options.

These examples show that the UIS RS already covers early learning to some extent in the reading and mathematics domains. If Task Force 4.2 support the use of the UIS RS in interim reporting, ACER-GEM can further develop these levels by drawing on additional assessments targeted specifically at the early years.

Describing a progression of holistic learning

Indicator 4.2.1 does not associate early learning with specific academic domains. This is an important difference between Indicator 4.1.1 and 4.2.1 that must be reflected in the UIS RS development. Instead, learning in 4.2.1 is a holistic concept reflecting the rapid onset of a wide range of capabilities, including those that influence the expression of a wide set of domain specific abilities (Anderson & Raikes, 2017). A key example is executive function – comprising capabilities such as self-regulation, attention, working memory –which plays an important role in enabling young children to engage in a wide range of interactions and environments that stimulate learning in a wide set of domains.

There is also strong evidence of a wider set of generalised set of cognitive abilities that precede domain-specific capabilities. For example, leveraging the Cattell–Horn–Carroll (CHC) model, cognitive abilities can be classified into domain-independent general capacities (eg, fluid reasoning and memory), acquired knowledge systems (eg, comprehension knowledge, reading, and writing), and sensory linked abilities (eg, visual and auditory processing). These are so-called meta-elements of cognitive functioning (Schneider & McGrew, 2012).

General abilities play an important role in the expression of academic capabilities (and not the other way around) (Deary, Strand, Smith, & Fernandes, 2007; Kaufman, Reynolds, Liu, Kaufman, & McGrew, 2012; Watkins, Lei, & Canivez, 2007). For example, short term memory and processing speed have been found to have significant effects on early reading (Floyd, Keith, Taub, & McGrew, 2007). These effects are stable over preschool and school ages, and the relationship between general abilities and domain-specific capabilities is especially relevant to the task of extending the UIS RS to the early years. This extension involves reconciling the existing domain-specific strands with a holistic definition of early learning.

Extending the UIS RS

The proposed approach to extending the UIS RS involves *retaining* the pre-reading and pre-mathematics skills described within the existing UIS RS domains, and *supplementing* these with a third "domain" representing general abilities. Compiling this third domain will involve extracting the general abilities already reflected in the domain-specific descriptions in the UIS RS.

This means that the UIS RS's description of early learning will offer more specific descriptions of learning progressions in:

- pre-reading
- pre-mathematics, and
- general abilities.

It is proposed that the general abilities domain may include several strands, based on theory and past empirical research: executive function, comprehension knowledge, fluid reasoning, long-term storage and retrieval, short-term memory/working memory. auditory processing, and verbal and non-verbal ability. These proposed strands will be refined as the domain is developed, to create a usable scale for 4.2.1 reporting.

Definition of the levels within these strands is already supported by the existing UIS RS descriptions. The skill illustrations and descriptions of strands describe various aspects of the tasks that require different levels of proficiency:

- the increasing cognitive abilities required; or example, the modality, length of the prompts children are required to attend to, and the number of things they need to remember

- the amount of competing information and its proximity to the relevant information is also described.

- the level of vocabulary (eg, comprehension knowledge/naming facility).

For example, the UIS RS skill illustrations in the retrieving strand at Levels 1 to 3 provide an example of a learning progression in self-regulation and short/term working memory. This can be seen in the increasing complexity of the location of the information, the amount of competing information, and the number if pieces of information that must be stored (all tasks are aural, supported with illustrations):

- Level 1: identify a (one) very simple piece of prominent information such as the main character's name

- Level 2: identify a few (one to three) prominent pieces of information such as a key event or the first, or last, event in a story with no competing information.

- Level 3: identify prominent pieces (one to three) of information from across the text with minimal competing information.

Similarly, in the UIS RS descriptions for mathematics, there are examples of non-verbal and verbal general abilities demonstrated by receptive language, vocabulary/naming facility, and the ability to understand directions (all tasks are to understand information presented verbally in a single short sentence):

- Level 1: understand the term 'smallest, and interpret related representations (groups of apples, cherries, chairs); find a single step strategy to identify the group with 'the smallest' and select their response from three or fewer options (point or say the name of the object).

- Level 2: understand the relational term 'most'; interpret quantities of a familiar real-world objects (groups of varying numbers of fruit) represented graphically in different arrangements; find a single step strategy to identify the group with 'the most' (eg, count and compare); and select their response from four options.

If Task Force 4.2 determines a need for additional descriptions of abilities commonly observed at lower levels, specifically those for children who would commonly be age 0-3, it is also possible to extend the UIS RS further, to the earliest foundations of learning and development. This can be done empirically by identifying exemplar items, adding them to the UIS RS, and building new descriptions of what would become Level 0 and lower.

Growth trajectories for the general cognitive abilities described above are observed across the life course. This conceptual model for extending the UIS RS is illustrated in Figure 1, showing the convergence of domain-specific and general abilities, and the potential for further extension of the scales to earlier age groups. The potential for extending the general abilities scale to higher levels is also shown, as a possible long-term aim that may be supported by recognition of these abilities in early learning.

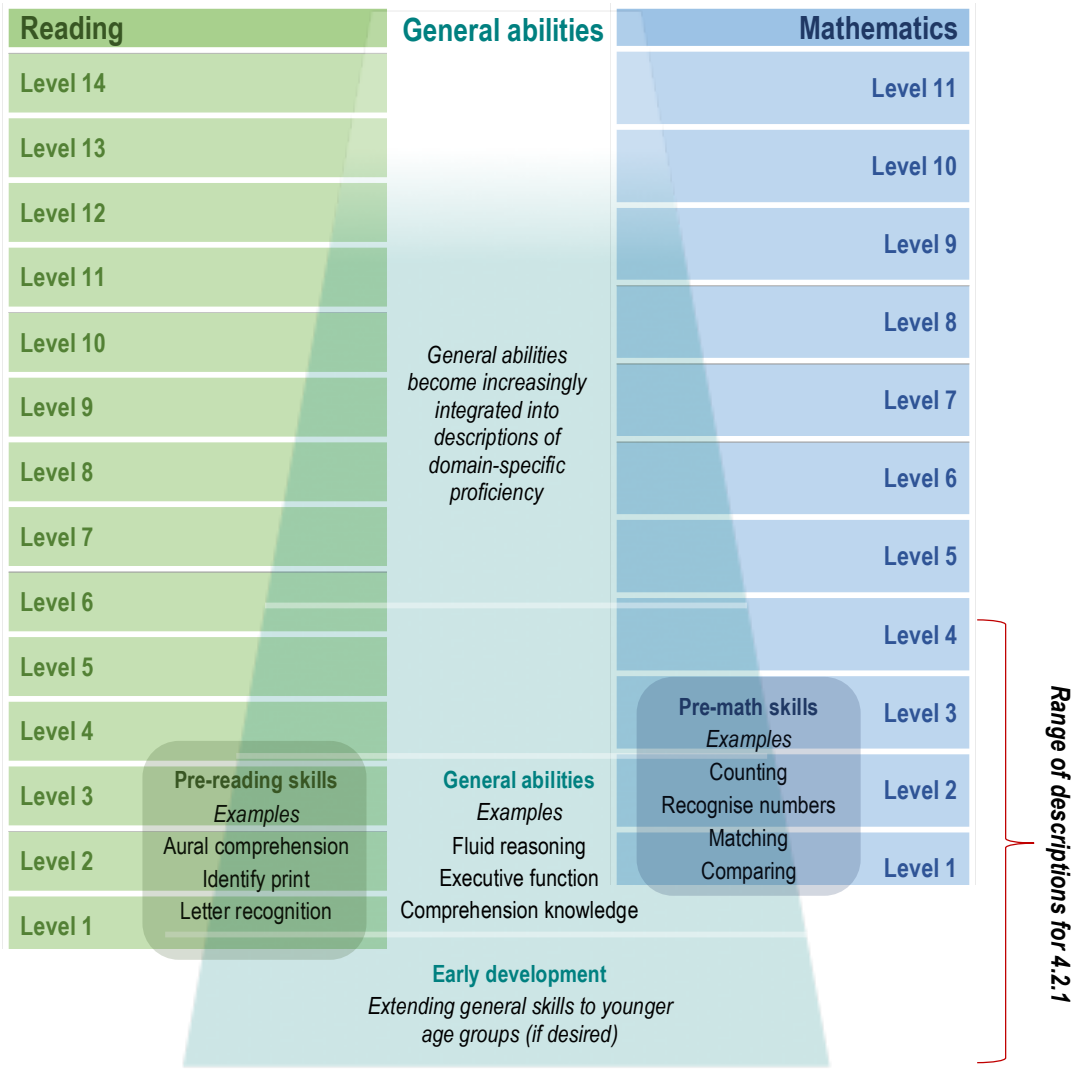

Figure 1: Framework for extension of UIS RS into *early learning*

Issues for consideration

The proposed approach is to provide as much flexibility as possible to ensure that many countries can report on learning for Indicator 4.2.1 whist maintaining a high level of quality and comparability. ACER-GEM recognises that there are many important issues to consider in any approach to the measurement of early learning, taking into account the needs of countries and the international assessment and reporting community, as well as children, families and early childhood development practitioners.

Some of these issues are outlined below.

Would it be better to have a single measure, used across all countries?

Using a common measure across all countries would provide the greatest consistency in 4.2.1 reporting – however, this is unlikely to occur in the near future. The UIS RS and Data Alignment process will help to improve interim reporting, for as long as multiple measures are used.

This does not preclude the adoption of a common measure (the ECDI) by countries over time. It is hoped that reporting using the UIS RS will help to generate appetite

for more consistent measures over time. Indeed, this process can include the recommendation that the easiest strategy for SDG 4.2.1 reporting is to adopt the ECDI in a national survey.

The UIS RS rests on the assumption that it is possible to create a common scale. What is the basis for this assumption?

The development of the UIS RS so far has involved detailed empirical work, to create a common learning progression (scale) from many assessments. The assessments were drawn from diverse contexts and include assessments conducted in developed and developing countries, and that target children and students from many ages and grades. This has supported the assumption that a common scale can be developed.

The next stage of UIS RS development involves testing this assumption in the field in 2018. ACER-GEM will undertake empirical validation of the UIS RS, and piloting and refining the Data Alignment process with a range of assessment programs. This will help to strengthen both the scales, and the process by which they may be used.

A similar process of conceptual development and validation would be necessary for the UIS RS in the early years. This would constitute a valuable exercise in itself, to test the consistency of early learning measures. As for 4.1.1, the quality and validity of the UIS RS will improve progressively over time, as this work progresses.

This validation process can be used to undertake the alignment of assessments to facilitate 4.2.1 reporting in 2018.

Does the extension of UIS RS into the early years imply the "schoolification" of early learning and assessment?

ACER-GEM is cognisant of the uniqueness of early childhood as an important stage of learning and development in its own right, as well as a precursor to later learning. The proposal to extend UIS RS to the early years does not imply any related extension of school-oriented curriculum, pedagogy or assessment practices. Indeed, the purpose of the UIS RS is to provide a non-prescriptive approach that enables a range of measures to be considered.

There remains a productive ongoing global debate about the best way of measuring early learning, in home-based or institutional settings (which is becoming increasingly possible, as participation in centre-based ECCE programs increases, particularly pre-primary education where significant increases in enrolment rate are observed even in developing contexts).

Work on extending the UIS RS into the early years, to align a variety of measures, can provide valuable information to inform this discussion.

Do the UIS RS domains (reading and mathematics) apply to early learning?

Different approaches are possible to extend the UIS RS domains into the early years. Our preferred approach outlined above supports a broader definition of early learning than domain-specific knowledge and skills alone; while also recognising that this broad base of learning includes knowledge and skills that are directly related to later reading and mathematics development.

What about other aspect of holistic learning?

The proposed strategy would define early learning under the three broad domains of general abilities, pre-reading, and pre-mathematics. There are other aspects of holistic learning not captured here, including for example social and emotional skills like empathy and trust, and motor development.

Certainly, some constraints must be placed around the definition of learning. Both for pragmatic reasons – there is only so much that can be assessed and reported on - but also because these aspects of learning (social skills, and motor) will also be measured as part of psychosocial wellbeing and health (the other aspects of 4.2.1 reporting that are outside the remit of this task force). That said, if there was

sufficient interest in broadening the reporting of psychosocial wellbeing and health too, similar methodologies as those used to develop the UIS RS could be undertaken within these domains.

Is learning the same in all contexts?

The 4.2 TF has identified reservations, for example, in the extent to which the development of decoding is invariant across all languages. Almost certainly there are examples of languages that differ significantly in the onset of fluency and mastery of decoding skills – some taking much longer that others, and being a relatively more difficult capability to demonstrate. There are empirical efforts underway now to quantify this effect within the scope of the UIS RS. That is, there is a pragmatic, or fit-for-purpose consideration to be made: some error is acceptable in order to facilitate international comparability.

In the meantime, it is plausible to take the position that, omitting decoding, other aspects of learning are invariant (or enough so for 4.2.1 reporting) that pre-reading comprehension skills, general abilities, and pre-mathematics can be used to describe a continuum of early learning across language contexts. This is consistent with international assessment programs and psychological measurement where assessments are equated across languages or empirically demonstrated to have acceptable between language variability.

Where to next?

The proposal in this paper provides the early stages of how the extension of the UIS RS into early learning may be conceptualised. If Task Force 4.2 endorse this approach in principle, further work is required to develop it more fully. This program of work could occur as part of the program of work that UIS and ACER-GEM is already pursuing, in refining and validating the UIS RS for 4.1.1 reporting.

To proceed with this work, ACER-GEM is seeking Task Force 4.2 endorsement of the following **recommendations**:

1. **That work on extending the UIS RS into early learning initially focus on skills that may be expected at ages 3–5, noting that there may be scope to develop the scales into earlier age ranges in future.**

2. **That this work include enhancing the Levels 1- to 3 in reading and mathematics to be suitable for this age group, as well as creating an additional domain to measure general abilities (see above).**

3. **That this work draw on existing measures of children's learning as its main empirical foundation, with a particular focus on alignment with the ECDI (but also other measures, including the MELQO MODEL, PRIDI, IDELA, EAP-ECDS).**

4. **That this work be guided by the need to identify a benchmark on the UIS RS for this age group, and reflect and inform any further Task Force 4.2 developments in defining *on track*.**

5. **That this work is framed primarily as supporting interim reporting against 4.2.1, noting that the development of ECDI remains the preferred longer-term strategy for consistent reporting.**

6. **That one outcome of this work is a series of case studies, drawing on the case studies developed for Data Alignment in 4.1.1, to illustrate how process can be used in practice. Case studies would leverage work done by Task Force members (Anderson & Raikes, 2017; Yoshikawa, Raikes, & Wuermli, 2017), and ACER-GEM would seek Task Force 4.2 input in their development and dissemination.**

7. **Ongoing consultation and input be sought from the TF 4.2 ECD experts, including Raikes, Yoshikawa, Janus, and Anderson as well as others as they wish to be included, during the implementation of interim reporting.**

References

ACER. (2017). Data Alignment for SDG 4 reporting A process to support the monitoring of learning outcomes for Sustainable Development Goal 4 Concept Note.

Anderson, K., & Raikes, A. (2017). Key Questions on the Domains of Measurement for SDG 4.2.1.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. Intelligence, 35, 13–21. https://doi.org/10.1016/j.intell.2006.02.001

Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell-Horn-Carroll Cognitive Abilities and Their Effects on Reading Decoding Skills: g Has Indirect Effects, More Specific Abilities Have Direct Effects. School Psychology Quarterly, 22(2), 200–233.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock–Johnson and Kaufman tests. Intelligence, 40, 123–138. https://doi.org/10.1016/j.intell.2012.01.009

Schneider, W., & McGrew, K. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (3rd ed., pp. 99–144). New York, NY: Guilford.

Watkins, M. W., Lei, P.-W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. Intelligence, 35, 59–68. https://doi.org/10.1016/j.intell.2006.04.005

Yoshikawa, H., Raikes, A., & Wuermli, A. (2017). Measurement Options for Development of Sustainable Development Goal Indicator 4.2.1 Memo, Global Alliance to Monitor Learning, Taskforce on Target 4.2.