

DESIGNING LEARNING ASSESSMENTS



HANDBOOK 3

Measuring Non-Academic Learning



Contents

Key terms	5
<hr/>	
1. Introduction	6
1.1 This handbook	7
<hr/>	
2. Defining and describing domains	8
2.1 Inter-Agency Network for Education in Emergencies (INEE) – competency types	8
2.2 Assessment & Teaching of 21 st Century Skills – ATC21S	9
2.3 Non-Academic Skills in Schools2030	10
<hr/>	
3. Validity and reliability in non-academic skills assessment	12
3.1 Validity risks in measuring non-academic skills	13
<hr/>	
4. Reviewing and adapting available tools	16
4.1 Reviewing available tools	16
4.2 Contextualising tools	20
<hr/>	
5. Developing and adapting instruments	22
5.1 Case study: Contextualising Psychosocial Adjustment in Sierra Leone	23
<hr/>	
Conclusion	25
<hr/>	
Example tools	26
International Social and Emotional Learning Assessment (ISELA)	26
Holistic Assessment of Learning and Development Outcomes (HALDO)	26
Social Emotional Assets and Resilience Scale (SEARS)	26
Measuring Early Learning Quality Outcomes (MELQO)	27
Explore SEL tool	27

Figures

Figure 1 INEE Competency types	8
Figure 2 ATC21S Skills	9
Figure 3 Non-academic domains for Schools2030	11
Figure 4 Steps involved in contextualising tools	20
Figure 5 Case study of contextualisation in Sierra Leone	25

Summary

This is the third in a series of handbooks on developing assessments. Handbook three is focused on the unique challenges of measuring non-academic skills. The content of the handbook covers four questions:

- 1

How do we define non-academic learning?

What frameworks have organisations used to define non-academic learning? How do they differ and what can we learn from them?

- The INEE Competency types framework.
 - The ACT21S framework for 21st Century Skills.
 - Framework for non-academic learning in Schools2030.
- 2

What are the risks to validity in measuring non-academic learning?

What are the specific challenges we face when measuring non-academic learning? What can we do to address some specific threats?

- Social desirability bias
 - Reference bias
 - Academic barriers to measuring non-academic learning
- 3

How do we review and adapt existing tools?

What kind of tools are we looking for? How do we contextualise them to the target learners?

- Reviewing available tools
 - Reflecting on kinds of questions and what will be most appropriate
 - The contextualisation process for non-academic measures
- 4

How do we develop new instruments?

What do we do if there is not one tool that we can adapt?

- Case study: contextualising Psychosocial adjustment in Sierra Leone

Key terms

TERM	DEFINITION
Social desirability bias	The tendency for respondents to under-report views and behaviours that are socially undesirable and over-report views and behaviours that are socially desirable.
Reference bias	The need for respondents to anchor their responses against a standard.
Anchoring vignettes	Hypothetical people and situations that provide a basis for students to provide a quantified response to a question.
Likert-type response	Participants select from a number (often 4 or 5) of ordered options, such as ‘strongly disagree’, ‘disagree’, ‘undecided’, ‘agree’ and ‘strongly agree’.
Likert scale	A scale produced from Likert-type responses.
Contextualisation	The process of adapting a tool to be appropriate for the intended population and context.
Think aloud protocol	A method of leading students to articulate their thought processes as they respond to the instrument so instrument developers can better understand how their questions work.



1 Introduction

This handbook is the third of a set of three designed to support learning assessment process in Schools2030. Together, the purpose of the handbooks is to provide a common understanding of the way that learning assessments work and how to conduct effective assessments.

Much of the content will be familiar to many, but the intention is that it can fill gaps or clarify ideas and processes for those who need it and provide grounds for common understanding across the programme. The handbooks are meant as a support and a guide, rather than an academic course.

The intended users are primarily the National Learning Partners, but all or parts of the handbooks can be shared with other Schools2030 partners and stakeholders to help explain elements of the assessment process.

The three handbooks are:



The core concepts handbook describes the core ideas that need to be considered throughout the process. It is theoretical in nature, but it is probably the most important of the three for those who are not familiar with measurement concepts.

The academic skills handbook walks through the process for developing assessments of academic skills. The non-academic skills handbook walks through the process for developing assessments for non-academic skills.

The three handbooks can be used together. The core concepts handbook should be the starting place for people who have limited familiarity with the concepts of validity, reliability and fairness. Others may want to read it quickly to make sure that any disagreement or confusion with these ideas can be clarified together at the outset.

The academic skills handbook is the main book to refer to while planning the design and implementation of an assessment of academic skills. Similarly, the non-academic skills handbook performs the same purpose for non-academic skills. They can then be referred to at each stage of the process. They are not comprehensive in detailing every consideration at every step, but they provide a basis for ensuring that important steps and considerations are not missed. At numerous points along the assessment process, validity, reliability and fairness will need to be discussed and considered. At these points it may be helpful to go back to the core concepts book to provide definitions to shape those discussions. Similarly, there may be parts of the academic and non-academic skills handbooks that prompt the reader to look back to the core concepts handbook.

It is intended that the handbooks can operate as reference materials so that they can be consulted as and when they are useful.

1.1 This handbook

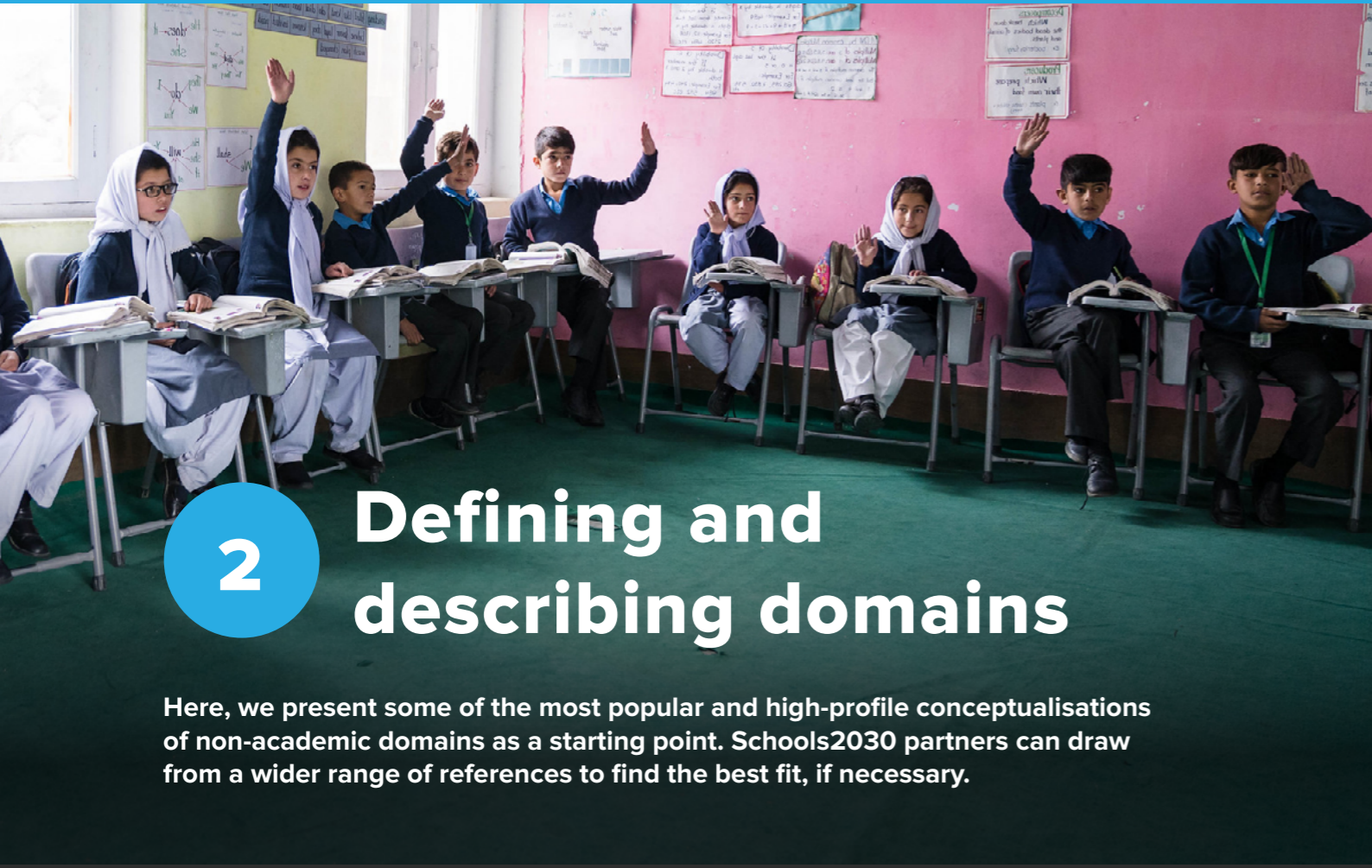
This handbook walks through the process of developing learning assessments of non-academic skills. It describes the main steps and discusses the main considerations. Throughout, validity, reliability and fairness are at the core of decision making. These are referred to regularly. Therefore, it is important that these concepts – discussed in the first handbook – are understood.

Section 2 discusses how some organisations have defined and described non-academic domains.

Section 3 then describes some threats to validity that are particularly problematic for non-academic assessments.

Section 4 walks through the process of reviewing and adapting existing instruments, which is likely to be the process followed for most (if not all) non-academic assessments in Schools2030.

Section 5 briefly discusses what is required to develop instruments without a pre-existing instrument upon which to base them. Finally, we present a case study, which illustrates the main steps of selecting and adapting instruments.



2 Defining and describing domains

Here, we present some of the most popular and high-profile conceptualisations of non-academic domains as a starting point. Schools2030 partners can draw from a wider range of references to find the best fit, if necessary.

2.1 Inter-Agency Network for Education in Emergencies (INEE) – competency types¹

The INEE divides competencies (including both academic and non-academic) into four types²:

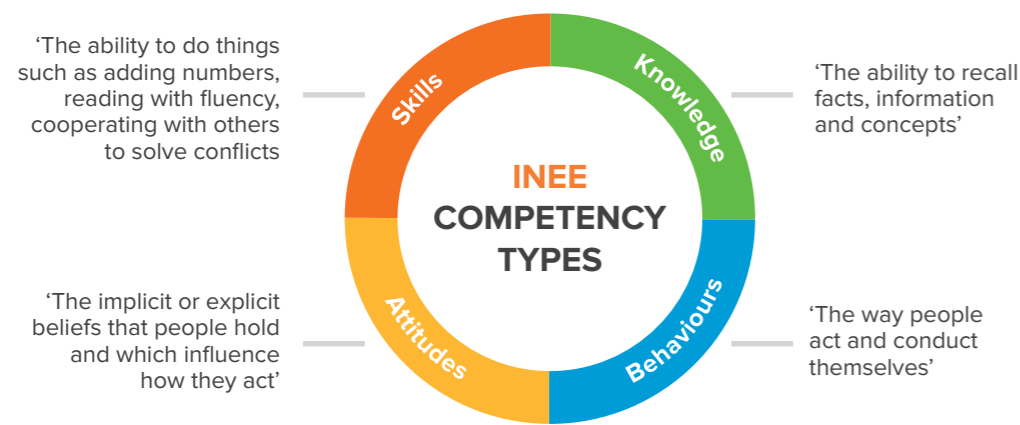


Figure 1 INEE Competency types

Competency domains can incorporate a set of knowledge, skills, attitudes and behaviours. For example, physical well-being may include health and hygiene, physical activity, sexual health and gross and fine motor skills. Socio-emotional processes may include emotion regulation, conflict resolution and cooperative behaviour.

The competency types provide a useful framework to help consider and define constructs.

2.2 Assessment & Teaching of 21st Century Skills – ATC21S³

ATC21S define ten 21st Century Skills and divide them into four categories shown in Figure 2. This framework shows how individual competencies can be mapped into categories, all of which relate to an outcome for a child, including academic outcomes, and life skills.



Figure 2 ATC21S Skills

1 - https://inee.org/system/files/resources/Measure_Guidance_final_11.19.2019.pdf
2- Ferrás, D. & Lee, J. Measure Guidance: Choosing and contextualizing assessment measures in educational contexts, 3EA.

3 - <https://resources.ats2020.eu/resource-details/LITR/ATC21s>

2.3 Non-Academic Skills in Schools2030

For Schools2030, 27 proficiencies and competencies have been defined (see Fig XX.). These are divided into four categories, aligned to the OECD’s skills for 2030 learning compass. The four areas are:

Knowledge

Knowledge encompasses the established facts, concepts, ideas and theories about certain aspects of the world. Knowledge usually includes theoretical concepts and ideas as well as practical understanding based on the experience of having performed certain tasks.⁴

Skills

Skills are the ability and capacity to carry out processes and to be able to use one’s knowledge in a responsible way to achieve a goal.⁵

Attitudes

Attitudes are underpinned by values and beliefs and have an influence on behaviour. Attitude reflects a disposition to react to something or someone positively or negatively and attitudes can vary according to specific contexts and situations.⁶

Values

Values are the guiding principles that underpin what people believe to be important when making decisions in all areas of private and public life. They determine what people will prioritise in making a judgement, and what they will strive for in seeking improvement.

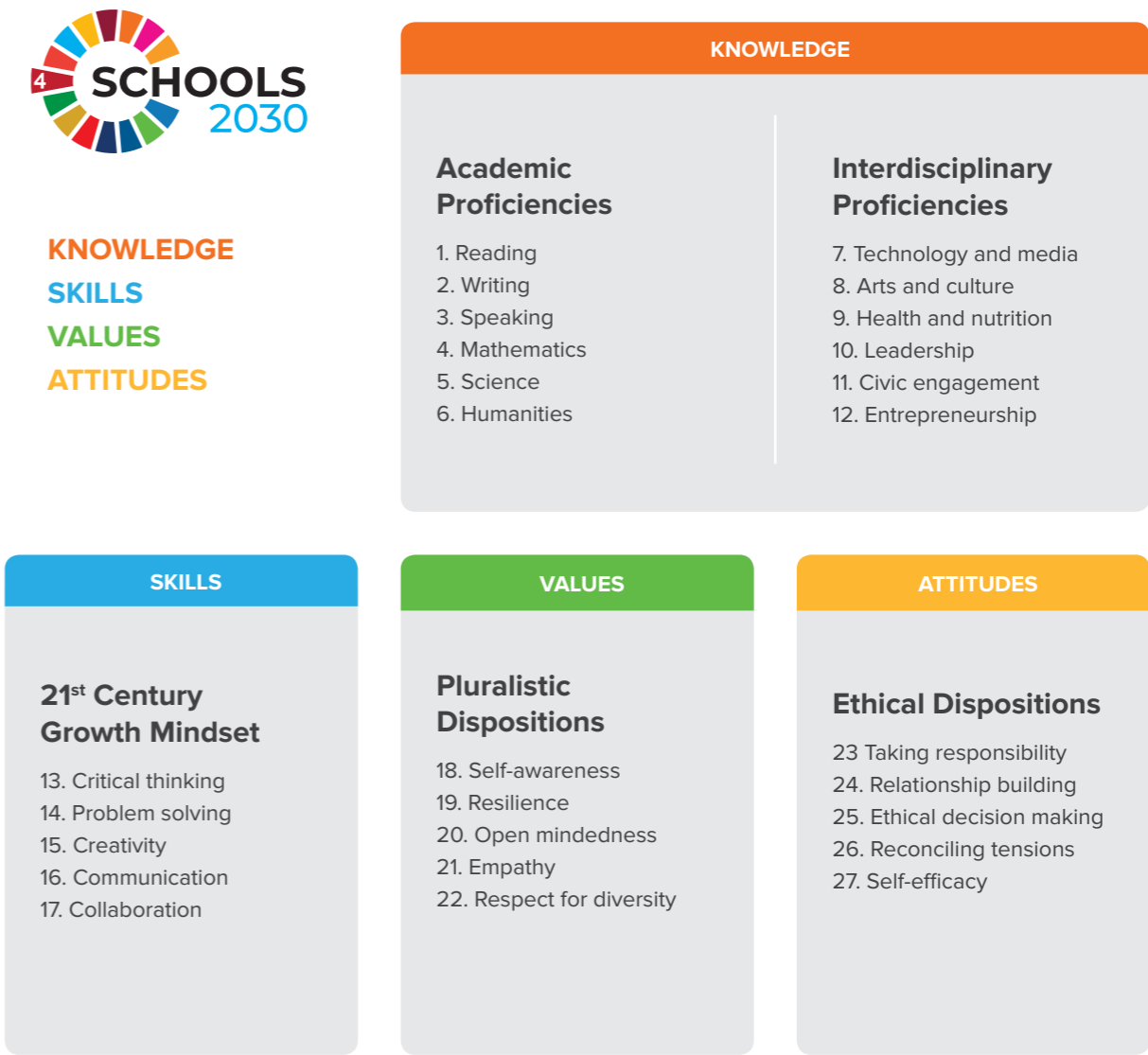


Figure 3 Non-academic domains for Schools2030

While these 27 domains were selected at the global level, the core of the selection process occurs at the national level. Schools2030 countries, as part of the assess phase, go through a rigorous and participatory process of selecting and defining domains of relevance for each country and age cohort. While the names of domains may be taken from the 27 suggested here, the definitions and constructs will be heavily influenced by content. This is crucial to ensure that what is measured reflects the understanding of and importance given to different domains in each country.

4 - See OECD concept note on Knowledge: https://www.oecd.org/education/2030-project/teaching-and-learning/learning/knowledge/Knowledge_for_2030_concept_note.pdf
5 - See OECD concept note on Skills: https://www.oecd.org/education/2030-project/teaching-and-learning/learning/skills/Skills_for_2030_concept_note.pdf
6 - See OECD concept note on Values and Attitudes: https://www.oecd.org/education/2030-project/teaching-and-learning/learning/attitudes-and-values/Attitudes_and_Values_for_2030_concept_note.pdf



3

Validity and reliability in non-academic skills assessment

The core handbook describes validity and reliability. The concepts apply equally to non-academic skills as to academic skills. Validity needs to be considered throughout the cycle.

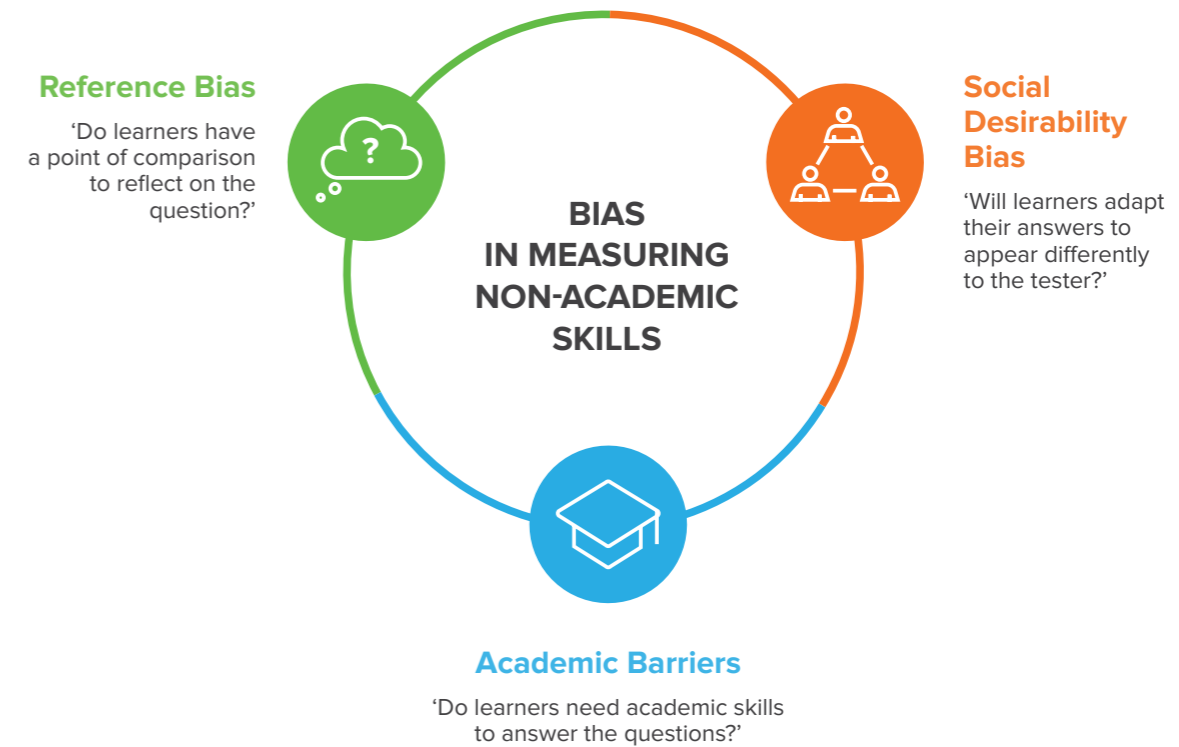
When we design assessments of non-academic skills, we face the same key challenge as we face when assessing academic skills: the measure often has to be narrower than the target construct. It is necessary to reduce the construct to something that is measurable for comparative purposes. There are added complications with many non-academic constructs because there can be a greater degree of subjectivity and variance in how skills are expressed. It is therefore particularly important to check throughout the process that the instrument assesses a construct that is defined objectively in a way that most people would agree with. The scoring should be designed so that the overall score is reflective of demonstrated attainment in the target construct. That is, you should expect that if people compared the levels of the target construct in two students using the questions or tasks used in the instrument, they should be able to draw a conclusion and their conclusion should be the same as that of the instrument, once the quantification and analysis has been completed.

Qualitative methods are important in the process of developing quantitative instruments to assess attainment in non-academic constructs. They can help to probe for the kinds of behaviours and responses that are representative of the construct within the context.

3.1 Validity risks in measuring non-academic skills

By nature, non-academic constructs can be even more sensitive to context than academic skills. They relate to behaviours and values that are very culturally specific, so it is important that instruments reflect the experiences of children. To do this, item and instrument development and adaptation must be driven by the context and not applied off the shelf.

A number of threats to validity can be particularly problematic for non-academic skills: social desirability bias, reference bias and academic barriers to expressing non-academic skills.⁷



Social desirability bias is the tendency for respondents to under-report views and behaviours that are socially undesirable and over-report views and behaviours that are socially desirable. This is a natural inclination and not any fault in respondents – we naturally want to be seen favourably. However, it does harm the validity of instruments.

7 - West et al (2015) Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling.

There are two key components: self-deception and impression management. Self-deception is the natural tendency for people to view themselves favourably. Impression management is actions taken as a result of peoples' desire to have other people view them positively.

IN ORDER TO MINIMISE DESIRABILITY BIAS, IT CAN BE HELPFUL TO:

ensure anonymity,

carefully consider who should administer the instrument,

carefully word items to avoid words that may trigger certain responses and

use indirect questioning so that the socially desirable response is less clear.



Finally, there may be **academic barriers to expressing non-academic skills**. Instruments need to be designed so that other factors – such as other skills and knowledge – do not mask the target construct. For example, poor numeracy and literacy skills could prevent a student from completing a creative problem-solving task even though they have the necessary problem-solving skills. For example, they may mis-understand the instructions or stimulus because of poor literacy or they may not have the arithmetic tools required to solve the problem. Instrument designers need to identify the skills and knowledge that are needed to answer the questions beyond the target construct. They need to design items that do not require additional skills that some test-takers do not possess. This can be summarised in two questions:

- What skills do children need in order to respond to the task correctly?
- Do all children being tested possess the necessary skills?

The answer to the second question must be ‘yes’.

This can be a significant challenge because many non-academic skills are closely related to academic skills. Academic skills and knowledge can provide the basis for non-academic skills. For example, when students are thinking critically or creatively, what are they thinking about? What tools do they use to think critically and what ideas and experiences do they depend on and combine when thinking creatively? For metacognition and awareness of their own learning and thinking, the learning and thinking will most often relate to academic constructs.

Similarly, reliability is as important to consider for non-academic assessment design as for academic assessment. Factors that may cause variation between groups or enumerators need to be identified and avoided. For example, it is important to keep consistent wording and prompting across all students because it is easy for enumerators to inadvertently prime students to give a certain response by saying things that are not in the script as an introduction to the next item.

Consider how students might be led to give certain answers and design the instrument and data collection methods to remain neutral and ensure that the data collected relates to the target construct and not to other characteristics of groups or inconsistencies in the implementation of the instrument.



Reference bias refers to the need for respondents to anchor their responses against a standard. For example, when asked about being hard working, they need to hold an idea of what that means in their mind. This standard may differ between different groups in different places meaning that the scores are not comparable.

Reference bias should be minimised by introducing clear reference points or by anchoring questions or using vignettes to make responses more comparable.

Anchoring vignettes help address reference bias. They describe hypothetical people and situations so that students can provide a quantified response. One example is that a student may consider where they sit on a scale in relation to the person described.



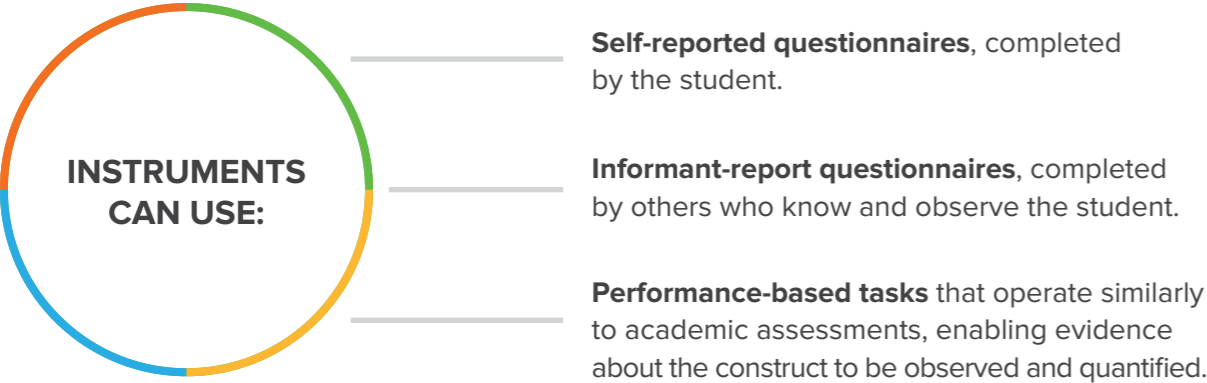
4 Reviewing and adapting available tools

Each has strengths and weaknesses, and some are better suited to different constructs and purposes from others. Performance-based tasks may provide better validity because they are less susceptible to some errors such as social desirability bias. However, it can be difficult to set meaningful tasks for some sub-domains and performance-based tasks are likely to be more complex and complicated to implement. The complexity also carries the risk of poorer inter-and intra-rater reliability as well as validity issues arising from difficulties in producing scores based on observed performance.

Reviewers should understand the benefits and disadvantages of each type of assessment in measuring the selected construct so that they can make rational and explicit decisions about which is more appropriate for the intended use.

4.1 Reviewing available tools

The first step should be identifying instruments that have been used to measure similar constructs. These instruments can be reviewed for possible items or whole tools that can be adapted and used. For most Schools2030 needs, this should provide a useful basis for the instrument. The Schools2030 community is also a useful resource for identifying instruments that others have used or considered.



Intra- and inter-rater reliability

- It is important to consider how scores from raters (e.g. teachers) may vary over time and how scores may differ between different raters.
- These inconsistencies in scoring are referred to as intra- and inter-rater reliability respectively.
- Intra-rater reliability is the degree to which scorers scores are consistent through time so that they reward the same score to answers of equal quality.
- Inter-rater reliability is the degree to which different scorers award the same scores to answers of equal quality.

Likert-type responses, Likert scales and factor analysis

It is likely that Likert scales will be used in instruments for non-academic constructs.

A **Likert-type response** is one where participants select from a number (often 5) of ordered options, such as ‘strongly disagree’, ‘disagree’, ‘undecided’, ‘agree’ and ‘strongly agree’. Likert-type responses are ordered, but the distance between each option is not constant or well-defined. The data is ordinal – it provides information about order (e.g. 5 is stronger agreement than 4, which is stronger agreement than 3 etc.). It is not interval because the difference between a 5 and a 4 is not precise and may not be the same as the difference between a 4 and a 3. Therefore data cannot analysed as if it were an interval scale.

A **Likert scale** is one constructed from Likert-type responses.

Factor analysis is useful to check that a group of Likert-type items combine to measure the same sub-domain. It describes the variability of responses and identifies the minimum number of unobserved variables (or factors) that may drive patterns of responses. If there are two factors, it means that a minimum of two unobserved variables could be driving results, based on observed correlations. For Likert scales, these factors are intended to be sub-domains, although factor analysis does not establish what the unobserved variables are. MeasurEd can provide support on analysis options once data is ready.

Batterton, M., Hale, K. (2017) The Likert Scale: What it is and how to use it. *Phalanx* 50(2)

The Empathy Quotient survey provides examples of Likert items. The respondents are requested to select one option for each statement: strongly disagree, slightly disagree, slightly agree or strongly agree.

	Strongly disagree	Slightly disagree	Slightly agree	Strongly agree
I can easily tell if someone else wants to enter a conversation.				
I find it difficult to explain to others things that I understand easily, when they don't understand it the first time.				
I really enjoy caring for other people.				
I try to solve my own problems rather than discussing them with others.				
I find it hard to know what to do in a social situation.				

When looking at available instruments to decide if they might work for your Schools2030 assessments, it is important to consider the target construct(s) as well as the context and student characteristics where the assessment will be administered.

Step 1: Record the constructs assessed for each option considered.

Step 2: Identify the commonalities and overlap with the target construct and how the construct measured by the instrument differs (significantly or subtly) from the domain.

Step 3: Map the constructs of the assessment instrument against the domain.

Step 4: Record the locations and populations that the instrument has been used for and consider how that might be different from the target for the Schools2030 assessment. Does the construct express differently in the places where the instrument has been used?

Step 5: Try to obtain psychometric data or reports about how the instrument performed from its previous uses. Was it valid and reliable? What evidence supports the validity and reliability claims? Were any issues identified? Where multiple options are available those with accompanying data and information should be preferred.

4.2 Contextualising tools

When an instrument or a combination of instruments have been identified to use as a starting point, the tools need to be contextualised.

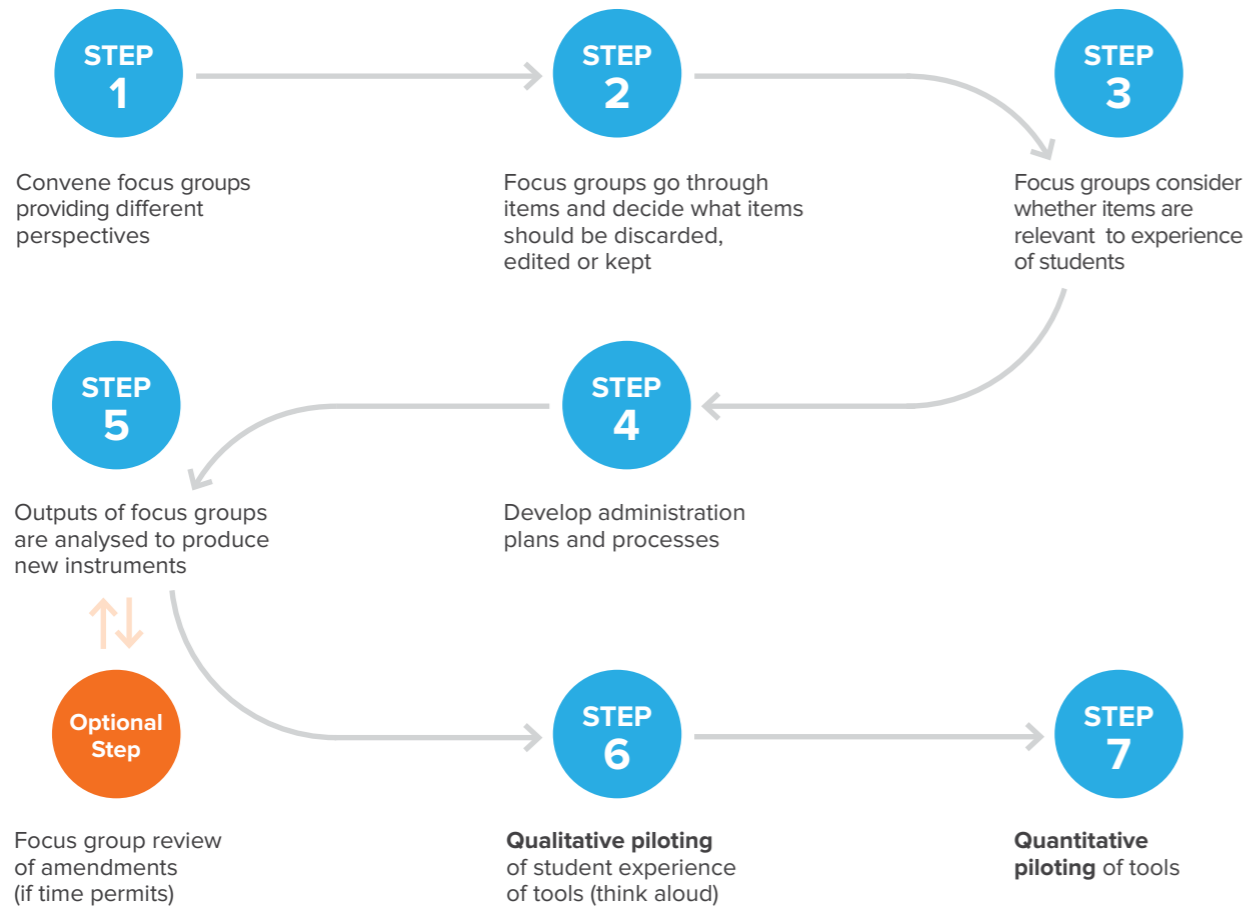


Figure 4 Steps involved in contextualising tools

The exact process can be designed to suit the purpose, context and resources, but it should entail some qualitative analysis of the instruments. The process below provides an example of how this might work.

Step 1: Focus groups providing different perspectives can be convened. Each should be familiar with the students and with the context. The focus groups first need to be introduced to the purpose of the assessment and the definition and conceptualisation of the construct. As views on these definitions often vary, it is likely to be productive to have a conversation about confusions and disagreements about the construct definition. Many Schools2030 country teams may have already gone through this process.

Step 2: In the focus groups, go through the source instruments and consider whether each item is applicable. Items can be discarded, edited or left as they are. The focus group discussion should also have time to consider whether the combination of remaining and edited items adequately cover the domain or whether there are significant gaps. They should also consider whether the balance is appropriate: Are all parts of the construct given the appropriate degree of prominence within the instrument?

Step 3: When considering the items, groups should particularly consider whether the items are relevant to the lived experiences of students. Are the examples and vignettes understandable and familiar to students? Are the expressions of the construct familiar to students? How do the focus groups expect students to feel and think as they respond to the questions?

Step 4: Important factors to consider when designing administration plans and processes will likely arise from consideration of how students might respond to questions and what factors might influence their responses. These should be noted and administration processes should be designed in parallel with the instrument.

Step 5: The results from focus groups then need to be analysed and aggregated to produce new instruments.

Optional Step: If time and resource permit and if significant changes have been made, it may be helpful to have further reviews (at a smaller level) of the amendments and check that the revised instrument is likely to work as intended.

Step 6: It may be helpful to use qualitative piloting methods for non-academic skills assessment. Given the greater risk that the instrument might not capture the intended construct, it can be helpful to add qualitative methods for some students to better understand how the students respond to the instrument. See the box on next page for an example of this.

Step 7: The quantitative piloting of instruments will be addressed in separate guidance, but it is an important step.

5

Developing and adapting instruments

Although it is expected that most, if not all, Schools2030 non-academic skills assessments will be based on pre-existing instruments, it is worth considering the process for developing instruments independently.

The process is not dissimilar to the adaptation process, but the starting point is further back. First there must be a qualitative investigation into how the construct is manifest within the target population. The starting point should still draw, as much as possible, from conceptualisations and definitions that have already been developed elsewhere. Focus groups should consider what this looks like for the target population.

From here, some draft instruments can be developed. Section 3 summarises the types of question – self-reports, informant-reports and performance based. Items should be mapped against the construct ensuring that it covers the range of skill levels that are likely to be observed in the target population. Anchoring vignettes are described in section 2. These may be helpful.

The process continues similarly to that described in section 3.2, but the degree of qualitative investigation into the instrument and piloting of the instrument will likely need to be greater to account for the fact that the instrument is new and untested.

5.1 Case study: Contextualising Psychosocial Adjustment in Sierra Leone

MacMullin and Loughry (2004)⁸ conducted a study into the psychosocial adjustment of former child soldiers in Sierra Leone and Uganda. They measured adjustments on four scales: anxiety and depression; pro-social behaviour; hostility and; confidence.

They based their instruments on two pre-existing tools that are built on commonly used child behaviour scales: the Child Behaviour Inventory (CBI)⁹ and the Cross-National Adolescent Project (C-NAP) questionnaire¹⁰. The CBI consist of 42 items that provide measures of depression, aggression, anxiety, pro-social behaviour and playful behaviour. It was designed for use in conflict-affected contexts and had been used in range of settings.

MacMullin and Loughry identified the main weaknesses of the CBI for their purposes as being its lack of specificity to the Sierra Leonean and Ugandan cultures and its inability to measure aspects of adjustment that may have been suggested by the target population. It also missed some of the sub-domains that they wanted to measure, namely self-concept and self-esteem.

The C-NAP contains 52 items that measure depression, aggression, delinquency, loneliness, social initiative, empathy and self-esteem. As with the CBI, the C-NAP instrument had been used in a range of countries including conflicted-affected settings. However, the C-NAP also shares the problem of being too generic.

Given the strengths of both instruments, the items from both were pooled to provide 94 items, which was reduced to 84 when duplicates were removed.

In Sierra Leone, MacMullin and Loughry consulted with six child service workers from the target area. They had planned to consult more widely, but security problems made that impossible. Three days of workshops took place, with the specific activities for each day shown in Figure 5.

In Sierra Leone, it was possible to include more people in the initial qualitative research stage. First, nine potential research assistants were trained. All had worked with youth for local or international Non-Governmental Organisations (NGOs). These nine research assistants undertook five days' training to equip them to assist with focus group discussions and to administer the final instrument. This group of nine formed the first focus group and then helped to facilitate three further focus group consultations. The focus group consultations followed a similar pattern to those in Uganda: (i) establishing a common understanding of the construct and purpose of the research; (ii) reviewing each item individually; (iii) identifying gaps and drafting

8 - MacMullin, C. and Loughry, M. (2004) Investigating psychosocial adjustment of former child soldiers in Sierra Leone and Uganda, *Journal of Refugee Studies* Vol 17(4).

9 - Macksoud et al. (1996) Assessing the impact of war on children, Apfel, R. and Simon, B. (eds) *Minefields in the heart*, New Haven, Yale University Press.

10 - Barber, B. (2001) *The Cross-National Adolescents Project*. University of Tennessee.

additional items.

Examples of the types of contextualised items that youth group members were able to generate include: “How often do you”:

- “Collect firewood, or join in other community activities?”
- “Have trouble in your heart?”
- “Sit around the fire at night?”
- “Think that you are not good at all?”

Items like this would not be found in generic instruments and provide clear information about adjustment as it is manifest in Sierra Leone.

DAY
1

- Explained the goals of the project and conducted a discussion about how former child soldiers adjusted and what indicators would be associated with the adjustments. This provided a solid understanding of the purpose and the construct.
- Items were considered one at a time and the group decided whether they thought each provided a good measure of child adjustment in the context. By the end of the day, 45 of the 84 items were identified as relevant for the Sierra Leonean measure.
- The 45 items were translated to Krio -the lingua franca, which all children would understand.
- The group were asked to identify additional indicators that were not included in the original list of 84.

DAY
2

- The Krio items were back translated into English. Discrepancies with the original English version were discussed and addressed by editing the Krio translation.
- A pilot instrument was constructed. The instrument asked children to decide how often they behaved, felt or thought in a particular way (e.g. “how often do you help the younger ones?”). Children selected responses from a 4-point Likert-type scale using the terms “never”, “rarely”, “sometimes” and “always”. Children could also choose “don’t know”.
- More modern approaches to translation in item development focus on translation for meaning into the target item and focus less on processes of translation and back translation to English.

DAY
3

- The researchers reported back on the instrument pilot.
- Based on the discussion of the pilot, four new items were drafted.

The instrument was then administered

Figure 5 Case study of contextualisation in Sierra Leone

Conclusion

This handbook has provided guidance for developing assessments of non-academic skills. Once the target constructs and assessment purpose are clearly defined, in most cases in Schools2030, the next step is to identify instruments that have been used for similar constructs. These can then be reviewed and adapted to produce instruments that provide the information that you require. It is important to stay mindful of the target construct and be aware of the risks to validity in this process.

There is a qualitative step in this development process that enables the instrument developers to better understand how the construct is expressed in their context and what the important nuances are. These lessons can then be built into a quantitative instrument.

The handbook is intended as a guide and we hope that it is a useful companion for the process. Oxford MeasurEd is available to provide support and the Assessment and Evaluation Hub is a useful starting place to discuss experiences and seek support.

Example tools

Below are some examples of robust instruments that have been used to assess non-academic skills. They act both as a reference point and a starting place when searching for available instruments.

International Social and Emotional Learning Assessment (ISELA)

ISELA assesses self-concept, stress management, perseverance, empathy, relationship management, and conflict resolution. It was designed for use with children aged 6 - 12 years old. It uses responses to scenarios and observed performance at some tasks.

➤ Available at:
<https://inee.org/measurement-library/international-social-and-emotional-learning-assessment-isela>

For more information, contact: learningassessment@savechildren.org.

Holistic Assessment of Learning and Development Outcomes (HALDO)

HALDO assesses both academic and non-academic constructs. It is designed to assess four domains: literacy, numeracy, social and emotional learning (SEL), and executive functioning. It is designed for use with children aged 4-12.

For more information, contact: learningassessment@savechildren.org

Social Emotional Assets and Resilience Scale (SEARS)

SEARS measures self-regulation, social competence, empathy and responsibility. It can be used for a wide range of ages. There are instruments that can be administered by parents or carers, teachers and children to provide different perspectives. They are based on observations or self-assessment, which are scored on a four-point scale.

➤ For more information, visit:
<https://www.parinc.com/Products/Pkey/406>

Measuring Early Learning Quality Outcomes (MELQO)

MELQO assesses working memory, inhibitory control, self-regulation, social cognition, social competence and emotional well-being. It is a freely-available measure developed by a number of multilateral organisations, predominantly the World Bank, UNICEF, UNESCO and the Brookings Institution. It is designed for use in early childhood education environments.

➤ For more information, visit:
<https://unesdoc.unesco.org/ark:/48223/pf0000248053>

Explore SEL tool

The Ecological Approaches to Social Emotional Learning (EASEL) Laboratory is hosted by Harvard University. It has produced an online tool for exploring SEL domains.

They have drawn together the available SEL frameworks to show how each conceptualises SEL and how the frameworks relate to each other.

Helpfully, where available, they link to assessment tools that relate to each framework. It takes some time to explore and see which frameworks are close to the construct definitions that you are using, but can be a helpful place to identify available tools.

➤ Access the tool at:
<http://exploresel.gse.harvard.edu>

DESIGNING LEARNING ASSESSMENTS

